

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/EP05/002557

International filing date: 10 March 2005 (10.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: DE
Number: 10 2004 014 274.2
Filing date: 22 March 2004 (22.03.2004)

Date of receipt at the International Bureau: 13 April 2005 (13.04.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

BUNDESREPUBLIK DEUTSCHLAND



PCT/EP200 5/ 0 0 2 5 57
18.03.05

Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 10 2004 014 274.2

Anmeldetag: 22. März 2004

Anmelder/Inhaber: Degussa AG, 40474 Düsseldorf/DE

Bezeichnung: Neue Alkoholdehydrogenasen

IPC: C 12 N 9/04

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 9. März 2005
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Schäfer

Neue Alkoholdehydrogenasen

Die Erfindung betrifft neue Polypeptide mit der biologischen Aktivität einer NAD- oder NADP-abhängigen Alkoholdehydrogenase. Des weiteren betrifft die Erfindung diese Polypeptide kodierende Nukleinsäuren, nicht-menschliche Wirte oder Wirtszellen sowie Reaktionssysteme, mit welchen gewünschte Produkte hergestellt werden können. Die erfindungsgemäßen Polypeptide werden bevorzugt in der Herstellung, ausgehend von Aldehyden bzw. Ketonen, von primären und enantiomerreinen, sekundären Alkoholen eingesetzt, die als Zwischenprodukte für Arzneimittel dienen können. Alternativ können die erfindungsgemäßen Polypeptide auch für die umgekehrte Reaktion, also die Oxidation von Alkoholen unter Ausbildung von Aldehyden bzw. Ketonen, eingesetzt werden.

In der Beschreibung ist auf eine Anzahl an Dokumenten verwiesen. Der Offenbarungsgehalt dieser Dokumente ist hiermit durch Bezugnahme inkorporiert.

Enantiomerenreine Alkohole gehören zu den bedeutendsten chiralen Bausteinen („Chiral Building Blocks“) der industriellen Spezial- und Feinchemie. Unter anderem fungieren diese Produkte als essentielle Schlüsselintermediate bei der Herstellung von Arzneimitteln. Der industrielle Weg zu diesen Zielmolekülen führte lange Zeit vorzugsweise über rein chemische Prozesse, beispielsweise auf dem Weg der Racematspaltung. Dabei wird ausgehend von einem Keton zunächst der Alkohol in racemischer Form hergestellt und dann das gewünschte Enantiomer in einer Racematspaltung unter Zuhilfenahme von mindestens stöchiometrischen Mengen eines chiralen Hilfsstoffes isoliert. Die Nachteile dieser Methodik sind nicht nur die maximal 50%-ige Ausbeute bei der Racematspaltung, sondern sind auch in der Verwendung zahlreicher ökologisch problematischer Einsatzstoffe bei der Herstellung des Racemats zu sehen. Weitere Nachteile bestehen in dem zusätzlichen Arbeitsschritt für das Recycling des ungewünschten Enantiomers, wie auch im Bedarf chiraler Hilfsreagenzien (noch dazu in stöchiometrischen Mengen) für die Racematspaltung. Das Konzept der Racematspaltung ist in Gleichung (1a) in der

Übersicht in Figur 1 zusammengefasst. Ein erster wesentlicher Fortschritt zu einem nachhaltigeren Verfahren konnte mit der biokatalytischen Racematspaltung erzielt werden, wodurch die Notwendigkeit des Einsatzes stöchiometrischer Mengen von chiralen Hilfsreagenzien entfällt. Bedauerlicherweise bleiben aber sämtliche anderen obig aufgezählten Nachteile trotz eines solchen biokatalytischen Weges relevant.

Ein Weg, bei dem vorstehend beschriebene Nachteile der Racematspaltung bzw. diastereoselektiver Synthesen vermieden werden können, besteht in der direkten Umwandlung von Ketonen in die gewünschten optisch aktiven Alkohole in einem Schritt. Solche „direkten asymmetrischen Prozesse“ können zum einen unter Verwendung von metallhaltigen Chemokatalysatoren durchgeführt werden, wobei als Chemokatalysatoren Schwermetall-haltige Komplexe, die einen chiralen Liganden beinhalten, zum Einsatz kommen. Neben dem Einsatz ökologisch problematischer Schwermetalle als wesentlicher Katalysatorkomponente ist auch der Bedarf an teuren und teilweise sehr empfindlichen Liganden, beispielsweise Phosphanliganden, nachteilig.

Eine weitere Alternative stellt die direkte asymmetrische Reduktion unter Verwendung geeigneter Biokatalysatoren für die quantitative Umwandlung von prochiralen Substraten zum gewünschten enantiomerenreinen Produkt dar. Die Anzahl der Reaktionsschritte wird auch hier auf das theoretisch mögliche Minimum von nur einem Schritt verkürzt, die biokatalytische Umwandlung erfolgt unter ökologisch vorzüglichen Bedingungen (u.a. Wasser als Solvens), und der Prozess als solcher verläuft mit hoher Atomökonomie. Das Konzept eines solchen biokatalytischen und nachhaltigen Verfahrens ist in Gleichung (1b) der Übersicht in Figur 1 aufgeführt.

Ein Nachteil der biokatalytischen Variante ist allerdings der Mangel an im industriellen Maßstab verfügbaren Alkoholdehydrogenasen als geeignete Biokatalysatoren für die Zielreaktion sowie deren Expression. Das der vorliegenden Erfindung zugrunde liegende technische Problem war deshalb, zu neuen, effizienten und industriell einsetzbaren Alkoholdehydrogenasen zu gelangen. Dieses technische Problem wird durch die in den Ansprüchen gekennzeichneten Ausführungsformen gelöst.

Somit betrifft die Erfindung ein Polypeptid mit der biologischen Aktivität einer NAD- oder NADP-abhängigen Alkoholdehydrogenase, das eine der nachfolgenden Sequenzen umfasst oder aufweist: die Sequenz der SEQ ID NO.: 1, die Sequenz der SEQ ID NO.: 2, die Sequenz der SEQ ID NO.: 3 oder eine Sequenz, die zu mindestens 90% identisch mit der Sequenz der SEQ ID NO.: 3 ist. Vorzugsweise ist die Sequenz zu mindestens 95% identisch mit der SEQ ID NO.: 3. Stärker bevorzugt ist die Sequenz zu mindestens 98% oder 99% identisch mit SEQ ID NO.: 3. Ferner umfasst ist die Sequenz der SEQ ID NO.: 4. Des weiteren umfasst ist die Sequenz der SEQ ID NO.: 5 oder eine Sequenz, die zu mindestens 90% identisch mit der Sequenz der SEQ ID NO.: 5 ist. Vorzugsweise ist die Sequenz zu mindestens 95% identisch mit der SEQ ID NO.: 5. Stärker bevorzugt ist die Sequenz zu mindestens 98% oder 99% identisch mit SEQ ID NO.:5. Ferner umfasst ist die Sequenz der SEQ ID NO.: 6 oder eine Sequenz, die zu mindestens 90% identisch mit der Sequenz der SEQ ID NO.: 6 ist. Vorzugsweise ist die Sequenz zu mindestens 95% identisch mit der SEQ ID NO.: 6. Stärker bevorzugt ist die Sequenz zu mindestens 98% oder 99% identisch mit der SEQ ID NO.: 6. Ebenfalls umfasst ist die Sequenz der SEQ ID NO.: 7 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 7 ist. Vorzugsweise ist die Sequenz zu mindestens 75% identisch mit der SEQ ID NO.: 7. Stärker bevorzugt ist die Sequenz zu mindestens 80% identisch mit der Sequenz der SEQ ID NO.: 7. Noch stärker bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 7. Ferner umfasst ist die Sequenz der SEQ ID NO.: 8 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 8 ist. Vorzugsweise ist die Sequenz zu mindestens 75% identisch mit der SEQ ID NO.: 8. Stärker bevorzugt ist die Sequenz zu mindestens 80% identisch mit der Sequenz der SEQ ID NO.: 8. Noch stärker bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 8. Darüber hinaus umfasst ist die Sequenz der SEQ ID NO.: 9 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 9 ist. Vorzugsweise ist die Sequenz zu mindestens 75% identisch mit der SEQ ID NO.: 9. Stärker bevorzugt ist die Sequenz zu mindestens 80% identisch mit der Sequenz der SEQ ID NO.: 9. Noch stärker bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 9. Ebenfalls umfasst ist die Sequenz der SEQ ID

NO.: 10 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 10 ist. Vorzugsweise ist die Sequenz zu mindestens 75% identisch mit der SEQ ID NO.: 10. Stärker bevorzugt ist die Sequenz zu mindestens 80% identisch mit der Sequenz der SEQ ID NO.: 10. Noch stärker bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 10. Ferner umfasst ist die Sequenz der SEQ ID NO.: 11 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 11 ist. Vorzugsweise ist die Sequenz zu mindestens 75% identisch mit der SEQ ID NO.: 11. Stärker bevorzugt ist die Sequenz zu mindestens 80% identisch mit der Sequenz der SEQ ID NO.: 11. Noch stärker bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 11. Des weiteren umfasst ist die Sequenz der SEQ ID NO.: 12 oder eine Sequenz, die zu mindestens 60% identisch mit der Sequenz der SEQ ID NO.: 12 ist. Vorzugsweise ist die Sequenz zu mindestens 65% identisch mit der SEQ ID NO.: 12. Stärker bevorzugt ist die Sequenz zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 12. Noch stärker bevorzugt ist die Sequenz zu mindestens 75% oder 80% identisch mit der SEQ ID NO.: 12. Besonders bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 12. Ebenfalls umfasst ist die Sequenz der SEQ ID NO.: 13 oder eine Sequenz, die zu mindestens 60% identisch mit der Sequenz der SEQ ID NO.: 13 ist. Vorzugsweise ist die Sequenz zu mindestens 65% identisch mit der SEQ ID NO.: 13. Stärker bevorzugt ist die Sequenz zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 13. Noch stärker bevorzugt ist die Sequenz zu mindestens 75% oder 80% identisch mit der SEQ ID NO.: 13. Besonders bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 13. Ferner umfasst ist die Sequenz der SEQ ID NO.: 14 oder eine Sequenz, die zu mindestens 75% identisch mit der Sequenz der SEQ ID NO.: 14 ist. Vorzugsweise ist die Sequenz zu mindestens 80% identisch mit der Sequenz der SEQ ID NO.: 14. Stärker bevorzugt ist die Sequenz zu mindestens 85% identisch mit der Sequenz der SEQ ID NO.: 14. Noch stärker bevorzugt ist die Sequenz zu mindestens 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 14. Darüber hinaus umfasst ist die Sequenz der SEQ ID NO.: 15 oder eine Sequenz, die zu mindestens 95% identisch mit der Sequenz der SEQ ID NO.: 15 ist. Vorzugsweise ist die Sequenz zu mindestens 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 15. Ferner

umfasst ist die Sequenz der SEQ ID NO.: 16 oder eine Sequenz, die zu mindestens 95% identisch mit der Sequenz der SEQ ID NO.: 16 ist. Vorzugsweise ist die Sequenz zu mindestens 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 16. Des weiteren umfasst ist die Sequenz der SEQ ID NO.: 17 oder eine Sequenz, die zu mindestens 75% identisch mit der Sequenz der SEQ ID NO.: 17 ist. Vorzugsweise ist die Sequenz zu mindestens 80% identisch mit der Sequenz der SEQ ID NO.: 17. Stärker bevorzugt ist die Sequenz zu mindestens 85% identisch mit der Sequenz der SEQ ID NO.: 17. Noch stärker bevorzugt ist die Sequenz zu mindestens 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 17. Ebenfalls umfasst ist die Sequenz der SEQ ID NO.: 18 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 18 ist. Vorzugsweise ist die Sequenz zu mindestens 75% identisch mit der SEQ ID NO.: 18. Stärker bevorzugt ist die Sequenz zu mindestens 80% identisch mit der Sequenz der SEQ ID NO.: 18. Noch stärker bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 18. Ferner umfasst ist die Sequenz der SEQ ID NO.: 19 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 19 ist. Vorzugsweise ist die Sequenz zu mindestens 75% identisch mit der SEQ ID NO.: 19. Stärker bevorzugt ist die Sequenz zu mindestens 80% identisch mit der Sequenz der SEQ ID NO.: 19. Noch stärker bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 19. Darüber hinaus umfasst ist die Sequenz der SEQ ID NO.: 20 oder eine Sequenz, die zu mindestens 60% identisch mit der Sequenz der SEQ ID NO.: 20 ist. Vorzugsweise ist die Sequenz zu mindestens 65% identisch mit der SEQ ID NO.: 20. Stärker bevorzugt ist die Sequenz zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 20. Noch stärker bevorzugt ist die Sequenz zu mindestens 75% oder 80% identisch mit der SEQ ID NO.: 20. Besonders bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 20. Ferner umfasst ist die Sequenz der SEQ ID NO.: 21 oder eine Sequenz, die zu mindestens 90% identisch mit der Sequenz der SEQ ID NO.: 21 ist. Vorzugsweise ist die Sequenz zu mindestens 95% identisch mit der Sequenz der SEQ ID NO.: 21. Stärker bevorzugt ist die Sequenz zu mindestens 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 21. Des weiteren umfasst ist die Sequenz der SEQ ID NO.: 22 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 22 ist.

Vorzugsweise ist die Sequenz zu mindestens 75% identisch mit der SEQ ID NO.: 22. Stärker bevorzugt ist die Sequenz zu mindestens 80% identisch mit der Sequenz der SEQ ID NO.: 22. Besonders bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 22. Ebenfalls umfasst ist die Sequenz der SEQ ID NO.: 23 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 23 ist. Vorzugsweise ist die Sequenz zu mindestens 60% identisch mit der SEQ ID NO.: 23. Stärker bevorzugt ist die Sequenz zu mindestens 65% identisch mit der Sequenz der SEQ ID NO.: 23. Noch stärker bevorzugt ist die Sequenz zu mindestens 70% oder 75% identisch mit der SEQ ID NO.: 23. Besonders bevorzugt ist die Sequenz zu mindestens 80%, 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 23. Ferner umfasst ist die Sequenz der SEQ ID NO.: 24 oder eine Sequenz, die zu mindestens 65% identisch mit der Sequenz der SEQ ID NO.: 24 ist. Vorzugsweise ist die Sequenz zu mindestens 70% identisch mit der SEQ ID NO.: 24. Stärker bevorzugt ist die Sequenz zu mindestens 75% identisch mit der Sequenz der SEQ ID NO.: 24. Noch stärker bevorzugt ist die Sequenz zu mindestens 80% oder 85% identisch mit der SEQ ID NO.: 24. Besonders bevorzugt ist die Sequenz zu mindestens 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 24. Des weiteren umfasst ist die Sequenz der SEQ ID NO.: 25 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 25 ist. Vorzugsweise ist die Sequenz zu mindestens 60% identisch mit der SEQ ID NO.: 25. Stärker bevorzugt ist die Sequenz zu mindestens 65% identisch mit der Sequenz der SEQ ID NO.: 25. Noch stärker bevorzugt ist die Sequenz zu mindestens 70% oder 75% identisch mit der SEQ ID NO.: 25. Besonders bevorzugt ist die Sequenz zu mindestens 80%, 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 25. Ferner umfasst ist die Sequenz der SEQ ID NO.: 26 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 26 ist. Vorzugsweise ist die Sequenz zu mindestens 60% identisch mit der SEQ ID NO.: 26. Stärker bevorzugt ist die Sequenz zu mindestens 65% identisch mit der Sequenz der SEQ ID NO.: 26. Noch stärker bevorzugt ist die Sequenz zu mindestens 70% oder 75% identisch mit der SEQ ID NO.: 26. Besonders bevorzugt ist die Sequenz zu mindestens 80%, 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 26. Ebenfalls umfasst ist die Sequenz der SEQ ID NO.: 27 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 27 ist. Vorzugsweise ist die

Sequenz zu mindestens 60% identisch mit der SEQ ID NO.: 27. Stärker bevorzugt ist die Sequenz zu mindestens 65% identisch mit der Sequenz der SEQ ID NO.: 27. Noch stärker bevorzugt ist die Sequenz zu mindestens 70% oder 75% identisch mit der SEQ ID NO.: 27. Besonders bevorzugt ist die Sequenz zu mindestens 80%, 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 27. Des weiteren umfasst ist die Sequenz der SEQ ID NO.: 28 oder eine Sequenz, die zu mindestens 75% identisch mit der Sequenz der SEQ ID NO.: 28 ist. Vorzugsweise ist die Sequenz zu mindestens 80% identisch mit der SEQ ID NO.: 28. Stärker bevorzugt ist die Sequenz zu mindestens 85% identisch mit der Sequenz der SEQ ID NO.: 28. Noch stärker bevorzugt ist die Sequenz zu mindestens 90%, 95%, 98% oder 99% identisch mit der SEQ ID NO.: 28. Ebenfalls umfasst ist die Sequenz der SEQ ID NO.: 29 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 29 ist. Vorzugsweise ist die Sequenz zu mindestens 60% identisch mit der SEQ ID NO.: 29. Stärker bevorzugt ist die Sequenz zu mindestens 65% identisch mit der Sequenz der SEQ ID NO.: 29. Noch stärker bevorzugt ist die Sequenz zu mindestens 70% oder 75% identisch mit der SEQ ID NO.: 29. Besonders bevorzugt ist die Sequenz zu mindestens 80%, 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 29. Ferner umfasst ist die Sequenz der SEQ ID NO.: 30 oder eine Sequenz, die zu mindestens 60% identisch mit der Sequenz der SEQ ID NO.: 30 ist. Vorzugsweise ist die Sequenz zu mindestens 65% identisch mit der SEQ ID NO.: 30. Stärker bevorzugt ist die Sequenz zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 30. Noch stärker bevorzugt ist die Sequenz zu mindestens 75% oder 80% identisch mit der SEQ ID NO.: 30. Besonders bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 30. Des weiteren umfasst ist die Sequenz der SEQ ID NO.: 31 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 31 ist. Vorzugsweise ist die Sequenz zu mindestens 60% identisch mit der SEQ ID NO.: 31. Stärker bevorzugt ist die Sequenz zu mindestens 65% oder 70% identisch mit der Sequenz der SEQ ID NO.: 31. Noch stärker bevorzugt ist die Sequenz zu mindestens 75% oder 80% identisch mit der SEQ ID NO.: 31. Besonders bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 31. Ebenfalls umfasst ist die Sequenz der SEQ ID NO.: 32 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 32 ist. Vorzugsweise ist die Sequenz zu mindestens 60% identisch mit

der SEQ ID NO.: 32. Stärker bevorzugt ist die Sequenz zu mindestens 65% oder 70% identisch mit der Sequenz der SEQ ID NO.: 32. Noch stärker bevorzugt ist die Sequenz zu mindestens 75 % oder 80% identisch mit der SEQ ID NO.: 32. Besonders bevorzugt ist die Sequenz zu mindestens 85%, 90%, 95%, 98% oder 99% identisch mit der Sequenz der SEQ ID NO.: 32. Des weiteren umfasst ist die Sequenz der SEQ ID NO.: 33 oder die Sequenz der SEQ ID NO.: 34. Die genannten und durch SEQ ID's gekennzeichneten Aminosäuresequenzen werden vorzugsweise durch die mit den SEQ ID Nummern 35 bis 68 bezeichneten DNA-Sequenzen kodiert. Bevorzugt sind ferner Polypeptide, die den natürlich vorkommenden Enzymen in voller Länge entsprechen. In einer anderen bevorzugten Ausführungsform enthalten die erfindungsgemäßen Polypeptide darüber hinaus mindestens einen heterologen Aminosäureabschnitt, der diese Polypeptide als Fusionsproteine kennzeichnet. Heterologe Bestandteile des erfindungsgemäßen Fusionsproteins können beispielsweise Tags (z. B. His-Tag oder Flag-Tag) sein, die bei der Aufreinigung der erfindungsgemäßen Fusionsproteine eingesetzt werden können. In anderen Ausführungsformen können die heterologen Bestandteile eine eigene enzymatische Aktivität aufweisen. In einem derartigen Fall sind die beiden enzymatischen Komponenten vorzugsweise durch einen Linker wie einen flexiblen 6-10 Aminosäure langen Glycin oder Glycin-Serin Linker verbunden, um die Funktionalität der Komponenten zu gewährleisten. Wie hier verwendet, kann der Begriff „heterolog“ einerseits bedeuten, dass die Komponenten des Fusionsproteins natürlicherweise nicht zusammen kovalent verbunden vorkommen, andererseits, dass die Komponenten aus verschiedenen Spezies stammen. Fusionsproteine werden üblicherweise durch rekombinante DNA-Technologie hergestellt (siehe Sambrook et al., a.a.O.).

Der Begriff „Polypeptid mit der biologischen Aktivität einer NAD- oder NADP-abhängigen Alkoholdehydrogenase“ bezeichnet erfindungsgemäß eine Gruppe von Enzymen, die die Umwandlung von Alkoholen in Aldehyde bzw. Ketone bzw. die entsprechende umgekehrte Reaktion, also die Umwandlung von Aldehyden in primäre Alkohole bzw. Ketone in sekundäre Alkohole, katalysiert. Die erstgenannte Reaktion entspricht dabei einem oxidativem Prozeß, der zweitgenannte Reaktionstyp ist ein reduktiver Prozeß. Die EC-Nummer von Alkoholdehydrogenasen (ADHs) ist EC 1.1.1.1. Neben den natürlicherweise auftretenden und im Zuge dieser Erfindung

isolierten Enzymen sind vom Schutzzumfang der Erfindung auch solche Polypeptide umfasst, die die vorstehend genannten Identitätswerte auf Aminosäureebene im Vergleich zu den aus natürlichen Quellen isolierten Polypeptiden aufweisen. Diese können ebenfalls aus natürlichen Quellen stammen. Andererseits können sie durch rekombinante DNA-Technologie dergestalt verändert werden, dass für den Fachmann vorhersehbar die enzymatische Aktivität erhalten oder im wesentlichen erhalten bleibt (vgl. beispielsweise Sambrook et al, „Molecular Cloning, A Laboratory Handbook“, 2. Auflage 1989, CSH Press, Cold Spring Harbor, Ausubel et al. „Current Protocols in Molecular Biology“, John Wiley & Sons, NY 2001). So können Aminosäuren, die sich nicht im aktiven Zentrum befinden und von denen prima facie nicht erwartet wird, dass der Austausch durch eine „gleichartige“ Aminosäure zu einer wesentlich veränderten dreidimensionalen Struktur führt, durch eine „gleichartige“ Aminosäure ausgetauscht werden. Beispielsweise kann erwartet werden, dass bestimmte Aminosäuren mit nicht polaren Seitenketten (gleichartige Aminosäuren), z.B. Alanin durch Valin, ausgetauscht werden können, ohne dass dies einen (wesentlichen) Einfluss auf die biologische Funktion des Enzyms, im Sinne der Erfindung auf die enzymatische Aktivität hätte. Auf der Basis seines Fachwissens kann der Fachmann entsprechende Schlussfolgerungen auch für den Austausch anderen Aminosäurearten (zum Beispiel den Ersatz basischer Aminosäuren durch andere basische Aminosäuren oder von Aminosäuren mit ungeladenen polaren Seitenketten durch andere Aminosäuren aus dieser Gruppe) erstellen.

Die prozentuale Identität zu den Aminosäuresequenzen der aus natürlichen Quellen isolierten Polypeptiden, die in dieser Beschreibung durch SEQ ID Nummern beschrieben sind, kann mit im Stand der Technik bekannten Verfahren vom Fachmann ohne weiteres ermittelt werden. Ein geeignetes Programm, das erfindungsgemäß eingesetzt werden kann, ist BLASTP (Altschul *et al.* 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(17):3389-3402.).

Ferner betrifft die Erfindung ein Nukleinsäuremolekül, das das erfindungsgemäße Polypeptid kodiert.

Das erfindungsgemäße Nukleinsäuremolekül kann ein DNA- oder ein RNA-Molekül sein. Bevorzugt ist, dass das Nukleinsäuremolekül ein cDNA- oder ein mRNA-Molekül ist. Erfindungsgemäß kann das DNA-Molekül des weiteren ein genomisches

DNA-Molekül sein. Ferner umfasst von der Erfindung sind Ausführungsformen, in denen das DNA-Molekül ein PNA-Molekül oder ein anderes Derivat eines DNA-Moleküls ist. Bevorzugt sind erfindungsgemäß DNA-Sequenzen, die die DNA-Sequenzen gemäß den SEQ ID Nummern 35 bis 68 umfassen.

Zur Lösung des der Erfindung zugrunde liegenden technischen Problems wurde der folgende Ansatz verfolgt. Auf Basis einer umfangreichen proprietären Stammsammlung erfolgte zunächst die Anzucht priorisierter Stämme auf Platte und nach erfolgter Viabilitäts- und Reinheitskontrolle auch in Flüssigkultur. Aus den geernteten Zellpellets wurde die genomische DNA dieser Organismen isoliert. Auf Basis der präparierten genomischen DNA wurde mittels erfindungsgemäßer Primer das genetische Screening ausgewählter Isolate auf Alkoholdehydrogenase-Gene via PCR-Typisierung durchgeführt. Auch die Homologie-bedingte Aminosäuresequenzähnlichkeit bereits bekannter Alkoholdehydrogenasen erlaubte dabei nicht ohne weiteres das Ableiten von Oligonukleotid-Primern, mit deren Hilfe bisher nicht identifizierte Alkoholdehydrogenase-Gene ohne weiteres erfolgreich amplifiziert werden können. Zunächst beruhte dieser Ansatz auf der Hypothese, dass bestimmte, in den bisher bekannten Alkoholdehydrogenase-Genen konservierte Sequenzmotive auch in den gesuchten neuen Alkoholdehydrogenase-Genen vorhanden sind. Die in den bisher bekannten Alkoholdehydrogenase-Genen konservierten Sequenzmotive sind jedoch ungeeignet, um über dem Fachmann bekannte Verfahren (Kwok *et al.*, 1995. Design and use of mismatched and degenerate primers. In: PCR Primer, A laboratory Manual, Dieffenbach CW & Dveksler GS (Editors), Cold Spring Harbor Laboratory Press, pp143-155; Compton T. 1990. Degenerate Primers for DNA Amplification. In: PCR Protocols, A Guide to Methods and Applications. Innis *et al.* (Editors) Academic Press, San Diego, pp39-34) degenerierte Primer abzuleiten. Die NAD- oder NADP-abhängigen Alkoholdehydrogenasen werden in langkettige, mittellange und kurzkettige ADHs eingeteilt. Sie werden vor allem auf Basis ihrer Metallabhängigkeit und der Größe der Untereinheiten in diese drei Gruppen eingeteilt. Die kurzkettigen ADHs benötigen keine Metall-Ionen und ihre Untereinheiten bestehen in etwa aus 250 Aminosäuren. Die mittellangen und langkettigen ADHs sind dagegen von Metall-Ionen abhängig. Die mittellangen, deren typische Untereinheiten aus ca. 350 Aminosäuren bestehen, benötigen Zink-Ionen. Die langkettigen ADHs, deren Untereinheiten aus ca. 385

Aminosäuren zusammengesetzt sind, benötigen Eisen-Ionen (Hummel, W. 1997. *New alcohol dehydrogenases for the synthesis of chiral compounds*. 58:145-84). Die Sequenzheterogenität nicht nur innerhalb aller bisher bekannten NAD- oder NADP-abhängigen ADHs ist ausgesprochen groß sondern auch die Sequenzheterogenität innerhalb der drei oben kurz beschriebenen ADH-Gruppen. Daher war es nicht ohne weiteres einfach, Primer zu konstruieren, mit Hilfe derer zum einen spezifisch ADH-Sequenzen amplifiziert werden können, und zum anderen auch eine zur Lösung der Aufgabe notwendige Diversität an neuartigen ADHs erfaßt werden kann. So wurden in den untersuchten Bakterien trotz des auf der Basis der verwendeten erwarteten Sequenzhomologien keinerlei neue langkettige ADHs isoliert. Die Qualität der konstruierten Primer sollte dabei zunächst mit genomischer DNA von Modellorganismen, deren Alkoholdehydrogenase-Gene bekannt sind oder mit DNA-Pools bestehend aus DNA aus verschiedenen Mikroorganismen ausgetestet werden. Dabei wurden PCR-Produkte kloniert, sequenziert und anschließend analysiert. Nach dieser Etablierungsphase wurde auf Basis der präparierten genomischen DNA der zu screenenden Mikroorganismen, wie oben beschrieben, eine PCR-Typisierung ausgewählter Isolate durchgeführt. Die aus den Experimenten gewonnenen Ergebnissen (zur Sequenzidentität und spezifischen Aktivität) flossen in die Priorisierung der potentiellen Hit-Organismen ein, deren neuartige Alkoholdehydrogenase-Gene isoliert werden. Trotz der unerwarteten, enttäuschenden und demotivierenden Ergebnissen bei der versuchten Isolierung von Nukleinsäuren, die langkettige ADHs kodieren sollten, wurde erfindungsgemäß eine Reihe an Nukleinsäuren isoliert, die kurzkettige Enzyme und mittellange Enzymketten kodieren. Diese wiesen z. T. eine erstaunlich niedrige Sequenzidentität (<50%) mit den bekannten Enzymen aus dieser Klasse auf.

Des weiteren betrifft die Erfindung ein Nukleinsäuremolekül, das komplementär zu dem erfindungsgemäßen Nukleinsäuremolekül ist.

Der Begriff „komplementär“ bedeutet erfindungsgemäß, dass sich die Komplementarität über den gesamten Bereich des erfindungsgemäßen Nukleinsäuremoleküls ohne Lücken erstreckt. Mit anderen Worten ist es erfindungsgemäß bevorzugt, dass die Komplementarität sich zu 100% über den gesamten Bereich der erfindungsgemäßen Sequenz, d.h. vom dargestellten 5'-Ende bis zum dargestellten 3'-Ende erstreckt. In weiteren bevorzugten Ausführungsformen

erstreckt sich die Komplementarität über einen Bereich von mindestens 19, bevorzugt mindestens 21 aufeinanderfolgenden Nukleotiden, die bevorzugt nicht für das aktive Zentrum der enzymatischen Aktivität kodieren.

Darüber hinaus betrifft die Erfindung einen Vektor, der das erfindungsgemäße Nukleinsäuremolekül enthält.

In den erfindungsgemäßen Vektoren liegen die erfindungsgemäßen Nukleinsäuren bevorzugt in operativer Verknüpfung mit einer Expressionskontrollsequenz vor, so dass sie in einer geeigneten Wirtszelle transkribiert und gegebenenfalls translatiert werden können. Expressionskontrollsequenzen umfassen üblicherweise einen Promotor und gegebenenfalls weitere regulatorische Sequenzen wie Operatoren oder Enhancer. Weiterhin können auch Translations-Initiationssequenzen vorhanden sein. Geeignete Expressionskontrollsequenzen für prokaryontische oder eukaryontische Wirtszellen sind dem Fachmann bekannt (siehe z.B. Sambrook et al., a.a.O.). Der erfindungsgemäße rekombinante Vektor kann weiterhin noch übliche Elemente wie einen Replikationsursprung und ein Selektionsmarkergen enthalten. Beispiele für geeignete rekombinante Vektoren sind Plasmide, Cosmide, Phagen, oder Viren (siehe z.B. Sambrook et al., supra). Ausgangsmaterialien für die Herstellung der erfindungsgemäßen rekombinanten Vektoren sind kommerziell erhältlich (z.B. von den Firmen Stratagene, InVitroGen oder Promega).

Alle dem Fachmann für diesen Zweck zur Verfügung stehenden Plasmide oder Vektoren kommen im Prinzip in Frage. Derartige Plasmide und Vektoren können z. B. von Studier und Mitarbeiter (Studier, W. F.; Rosenberg A. H.; Dunn J. J.; Dubendorff J. W.; (1990), Use of the T7 RNA polymerase to direct expression of cloned genes, *Methods Enzymol.* 185, 61-89) oder den Broschüren der Firmen Novagen, Promega, New England Biolabs, Clontech oder Gibco BRL entnommen werden. Weiter bevorzugte Plasmide und Vektoren können gefunden werden in: Glover, D. M. (1985), *DNA cloning: a practical approach*, Vol. I-III, IRL Press Ltd., Oxford; Rodriguez, R.L. und Denhardt, D. T (eds) (1988), *Vectors: a survey of molecular cloning vectors and their uses*, 179-204, Butterworth, Stoneham; Goeddel, D. V. (1990), *Systems for heterologous gene expression*, *Methods Enzymol.* 185, 3-7; Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York. Plasmide, mit denen das die erfindungsgemäße Nukleinsäure aufweisende

Genkonstrukt in ganz bevorzugter Weise in den Wirtsorganismus kloniert werden kann, sind Derivate von: pUC18 und pUC19 (Roche Biochemicals), pKK-177-3H (Roche Biochemicals), pBTac2 (Roche Biochemicals), pKK223-3 (Amersham Pharmacia Biotech), pKK-233-3 (Stratagene) oder pET (Novagen). Weitere bevorzugte Plasmide sind pBR322 (DSM3879), pACYC184 (DSM4439) und pSC101 (DSM6202), welche von der DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany bezogen werden können. Bevorzugte Promotoren sind beispielsweise T7, lac, tac, trp, rha und ara.

Ferner betrifft die Erfindung einen nicht-menschlichen Wirt, der das erfindungsgemäße Polypeptid oder das erfindungsgemäße Nukleinsäuremolekül oder den erfindungsgemäßen Vektor enthält.

Der nicht-menschliche Wirt kann eine Zelle oder ein mehr- bis vielzelliger Organismus sein. Geeignete vielzellige Organismen schließen in der Molekularbiologie geläufige Modellsysteme wie *Drosophila melanogaster*, Zebrafisch oder *C. elegans* ein.

In einer bevorzugten Ausführungsform ist der Wirt eine Zelle.

Der erfindungsgemäße Wirt ist in dieser bevorzugten Ausführungsform eine rekombinante Zelle, die mit einer erfindungsgemäßen Nukleinsäure oder einem erfindungsgemäßen Vektor transformiert oder transfiziert wurde (die Begriffe „Transformation“ und „Transfektion“ werden gemäß dieser Erfindung sinngleich verwendet). Die Transformation bzw. Transfektion kann nach bekannten Methoden erfolgen, z.B. durch Calciumphosphat-Copräzipitation, Lipofektion, Elektroporation, Partikelbeschuss oder virale Infektion. Die erfindungsgemäße Zelle kann die rekombinante Nukleinsäure in extrachromosomaler oder chromosomal integrierter Form enthalten. Mit anderen Worten kann die Transfektion/Transformation eine stabile oder transiente Transfektion/Transformation sein.

Vorzugsweise ist die rekombinante Zelle prokaryotischen Ursprungs. Geeignete Wirtszellen schließen Zellen von einzelligen Mikroorganismen wie bakterielle Zellen ein. Als bakterielles Wirtssystem eignet sich insbesondere *E. coli*. *E. coli* weist im Zytoplasma die Cofaktoren auf, die für die enzymatische Aktivität des erfindungsgemäßen Polypeptids benötigt werden. Dies sind insbesondere NADH, NADPH, NAD⁺ oder NADP⁺. Ganz besonders bevorzugt sind: *E. coli* XL1 Blue,

W3110, DSM14459 (PCT/US00/08159), NM 522, JM101, JM109, JM105, RR1, DH5, TOP 10- oder HB101. Ferner können Bakterien der Genera/Spezies *Lactobacillus*, *Bacillus*, *Rhodococcus*, *Campylobacter*, *Caulobacter*, *Mycobacterium*, *Streptomyces*, *Neisseria*, *Ralstonia*, *Pseudomonas*, sowie *Agrobacterium* für die Expression der erfindungsgemäßen Nukleinsäuren eingesetzt werden. Entsprechende Stämme sind im Stand der Technik verfügbar und können, zumindest teilweise, über die internationalen Hinterlegungsstellen wie die ATCC oder die DMSZ bezogen werden. Transfektions- und Transformations-Protokolle sind dem Fachmann bekannt (Chan and Cohen. 1979. High Frequency Transformation of *Bacillus subtilis* Protoplasts by Plasmid DNA. *Mol Gen Genet.* 168(1):111-5; Kieser *et al.*. 2000. *Practical Streptomyces Genetics*. The John Innes Foundation Norwich.; Sambrook *et al.*. 1989. *Molecular Cloning. A Laboratory Manual*. In: second ed.. Cold Spring Harbor Laboratory Press. Cold Spring Harbor. NY.; Irani and Rowe. 1997. Enhancement of transformation in *Pseudomonas aeruginosa* PAO1 by Mg^{2+} and heat. *Biotechniques* 22: 54-56). Darüberhinaus können als Wirtsorganismen beispielsweise auch Hefen wie *Hansenula polymorpha*, *Pichia sp.*, *Saccharomyces cerevisiae* herangezogen werden.

Alternativ dazu kann die Zelle eukaryontischen Ursprungs sein. Geeignete eukaryontische Zellen schließen CHO-Zellen, HeLa-Zellen und andere ein. Viele dieser Zellen sind über Hinterlegungsstellen wie die ATCC oder die DMSZ erhältlich.

In einer weiteren bevorzugten Ausführungsform ist der Wirt ein transgenes nicht-menschliches Tier.

Transgene nicht-menschliche Tiere können nach im Stand der Technik bekannten Verfahren hergestellt werden.

Das erfindungsgemäße transgene nicht-menschliche Tier kann bevorzugt verschiedene genetische Konstitutionen aufweisen. Es kann (i) das Gen einer erfindungsgemäßen Nukleinsäure konstitutiv oder induzierbar überexprimieren, (ii) das endogene Gen einer erfindungsgemäßen Nukleinsäure in inaktivierter Form enthalten, (iii) das endogene Gen einer erfindungsgemäßen Nukleinsäure vollständig oder teilweise durch ein mutiertes Gen einer erfindungsgemäßen Nukleinsäure ersetzt enthalten, (iv) eine konditionale und gewebsspezifische Überexpression oder

Unterexpression des Gens einer erfindungsgemäßen Nukleinsäure aufweisen oder (v) einen konditionalen und gewebsspezifischen Knock-out des Gens einer erfindungsgemäßen Nukleinsäure aufweisen.

Vorzugsweise enthält das transgene Tier zusätzlich ein exogenes Gen einer erfindungsgemäßen Nukleinsäure unter Kontrolle eines die Überexpression erlaubenden Promotors. Alternativ kann das endogene Gen einer erfindungsgemäßen Nukleinsäure durch Aktivierung oder/und Austausch des eigenen Promotors überexprimiert werden. Vorzugsweise weist der endogene Promotor des Gens einer erfindungsgemäßen Nukleinsäure eine genetische Veränderung auf, die zu einer erhöhten Expression des Gens führt. Die genetische Veränderung des endogenen Promotors umfasst dabei sowohl eine Mutation einzelner Basen als auch Deletions- und Insertionsmutationen.

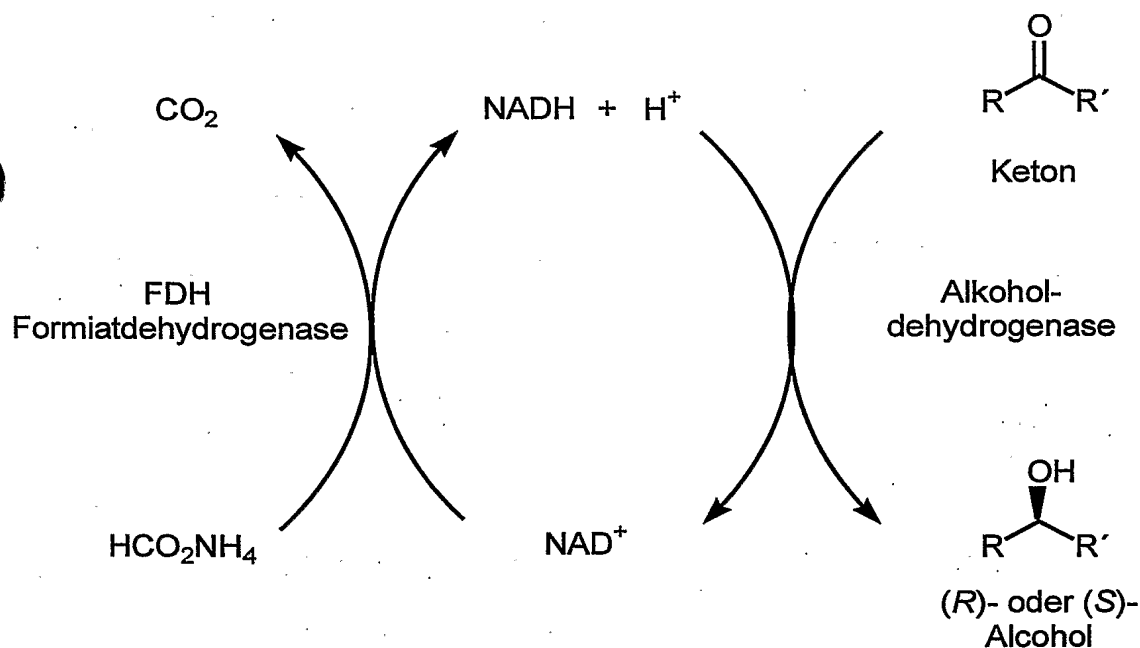
In einer besonders bevorzugten Ausführungsform des erfindungsgemäßen Wirts ist dieser ein transgener Nager, vorzugsweise eine transgene Maus, ein transgenes Kaninchen, eine transgene Ratte, oder ein transgenes Schaf, eine transgene Kuh, eine transgene Ziege oder ein transgenes Schwein.

Mäuse haben gegenüber anderen Tieren zahlreiche Vorteile. Sie sind leicht zu halten und ihre Physiologie gilt als Modellsystem für die des Menschen. Die Herstellung solcher Gen-manipulierten Tiere ist dem Fachmann hinreichend bekannt und wird nach üblichen Verfahren durchgeführt (sh. z.B. Hogan, B., Beddington, R., Costantini, F. und Lacy, E. (1994), *Manipulating the Mouse-Embryo; A Laboratory Manual*, 2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; WO91/08216).

Alternativ oder zusätzlich können auch Zellkultursysteme, insbesondere humane Zellkultursysteme, für die Anwendungen eingesetzt werden, die für das erfindungsgemäße nicht-menschliche transgene Tier beschrieben sind.

Als Cofaktoren der Alkoholdehydrogenasen werden – wie bereits vorstehend erwähnt, in Abhängigkeit von der jeweiligen Alkoholdehydrogenase - beispielsweise NADH bzw. NADPH bzw. deren oxidierte Formen NAD^+ und NADP^+ eingesetzt.

Die Regeneration der Cofaktoren kann im Prinzip entweder enzymgekoppelt unter Einsatz eines zweiten Enzyms, beispielsweise einer Formiatdehydrogenase oder einer Glucosedehydrogenase, oder substratgekoppelt unter Einsatz eines von der eingesetzten Alkoholdehydrogenase als Substrat akzeptierten Alkohols, beispielsweise – falls als Substrat akzeptiert – *iso*-Propanol, erfolgen. Exemplarisch ist im nachfolgenden Schema das Konzept der Alkoholdehydrogenase-katalysierten Reduktion eines Ketons mit einer enzymgekoppelten Cofaktorregenerierung unter Einsatz einer Formiatdehydrogenase aufgezeigt.



Alkoholdehydrogenasen werden beispielsweise zur Herstellung von enantiomerenangereicherten, vorzugsweise enantiomerenreinen sekundären Alkoholen ausgehend von prochiralen Ketonen eingesetzt. Dabei sind sowohl (R)- als auch (S)-spezifische Alkoholdehydrogenasen bekannt, die entsprechend zur Ausbildung der jeweiligen enantiomeren (R)- bzw. der (S)-Form von Alkoholen führen.

Dementsprechend ist erfindungsgemäß auch ein Wirt bevorzugt, der eine weitere, für die Cofaktor-Regenerierung geeignete Dehydrogenase oder ein diese Dehydrogenase kodierendes Nukleinsäuremolekül aufweist.

Dabei kann der Wirt diese weitere Dehydrogenase natürlicherweise enthalten oder er kann mit einer rekombinanten Nukleinsäure, die diese Dehydrogenase kodiert und mit der die Dehydrogenase in dem Wirt exprimiert werden kann, transfiziert werden

sein. Diese Ausführungsform setzt auch voraus, dass der Wirt die für die Funktion dieser weiteren Dehydrogenase notwendigen Cofaktoren aufweist oder dass ihm diese Cofaktoren in geeigneter Weise zugeführt werden.

Insbesondere bevorzugt ist in diesem Zusammenhang ein Wirt, wobei die für die Cofaktor-Regenerierung geeignete Dehydrogenase eine Formiatdehydrogenase oder eine Glucosedehydrogenase ist. Besonders bevorzugt ist eine Formiatdehydrogenase aus *Candida boidinii*. Ferner ist besonders bevorzugt, dass die cofaktorregenerierende Dehydrogenase eine Glucosedehydrogenase aus *Bacillus subtilis* ist. Gentechnisch veränderte Mutanten der genannten cofaktorregenerierenden Dehydrogenasen, die die genannte enzymatische Funktion beibehalten, sind erfindungsgemäß ebenfalls bevorzugt.

In einer weiteren Ausführungsform betrifft die Erfindung ein Reaktionssystem, das eine ein Substrat einer Dehydrogenase darstellende organische Verbindung enthält, ferner das erfindungsgemäße Polypeptid, den erfindungsgemäßen Vektor oder den erfindungsgemäßen Wirt und gegebenenfalls einen Cofaktor für das erfindungsgemäße Polypeptid. (Cofaktor-Zugabe ist in den Fällen nötig, in denen der Cofaktor nicht schon im System vorhanden ist, siehe auch unten) Das erfindungsgemäße Reaktionssystem kann in einem Falle eine bakterielle Zelle sein, die dem erfindungsgemäßen Wirt entspricht und das erfindungsgemäße Polypeptid sowie die notwendigen Cofaktoren im Zytoplasma aufweist. In dem Falle, dass der/die Cofaktor(en) bereits natürlicherweise im System/Wirt vorhanden sind, müssen sie nicht mehr gesondert zugeführt werden. Geeigneterweise ist der Wirt ein solcher, der eine weitere, für die Cofaktor-Regenerierung geeignete Dehydrogenase oder ein diese Dehydrogenase kodierendes Nukleinsäuremolekül aufweist und ferner die dazu notwendigen Cofaktoren. Wenn diesem Reaktionssystem ein Substrat für ein gewünschtes Produkt zugeführt wird oder dieses im Reaktionssystem selbst metabolisiert wird, so kann das gewünschte Produkt ohne weiteres aus dem Reaktionssystem isoliert werden, wenn das Reaktionssystem unter geeigneten Bedingungen gehalten wird. Geeignete Bedingungen schließen ein, dass die Reaktion bei Temperaturen von 10 bis 80°C, bevorzugt von 20 bis 60°C, und ganz bevorzugt von 20 bis 40°C durchgeführt wird. Bevorzugt ist auch, dass die Substratkonzentration bei 100 bis 2000 mM, vorzugsweise bei 200 bis 800 mM liegt.

In bevorzugter Form wird die gewünschte Reaktion so durchgeführt, dass Umsätze von $>80\%$, insbesondere $>90\%$, innerhalb von <20 Stunden Reaktionszeit, insbesondere <10 Stunden Reaktionszeit und ganz bevorzugt <5 Stunden Reaktionszeit erreicht werden. Das Reaktionssystem kann in einer anderen Ausführungsform ein in vitro System zur Umsetzung eines geeigneten Substrats zum Erhalt des gewünschten Produktes sein. Beispielsweise kann das erfindungsgemäße Polypeptid mit den genannten Cofaktoren und dem Substrat und gegebenenfalls (d.h. sofern nötig) mit einer weiteren, für die Cofaktor-Regenerierung geeigneten Dehydrogenase (und gegebenenfalls dafür notwendigen Cofaktoren, insbesondere NADH und NADPH bzw. deren oxidierte Formen) unter geeigneten Bedingungen, wie beispielsweise vorstehend ausgeführt und für einen ausreichenden Zeitraum in Kontakt gebracht werden, so dass das gewünschte Produkt entstehen kann. Bei dieser in vitro-Variante unter Nutzung von isolierten Enzymen (in gereinigter Form oder als Rohextrakt) und Zusatz von Cofaktoren sollten diese Cofaktorzusätze im Sinne einer ökonomischen Prozessführung bei <0.01 Äquivalenten (bezogen auf die eingesetzte Substratmenge), bevorzugt <0.001 Äquivalenten und ganz bevorzugt <0.0005 Äquivalenten liegen.

Ferner kann das „Reaktionssystem“ auch ein transgenes nicht-menschliches Tier sein, dem ein geeignetes Substrat und gegebenenfalls Cofaktoren oder/und die weitere Dehydrogenase gefüttert oder verabreicht wird und das das Substrat in geeigneten Geweben umsetzen kann. Das Reaktionssystem kann in einer anderen Ausführungsform auch ein zelluläres Membransystem sein, in dem das Enzym, die Enzyme und ggf. die Cofaktoren verankert sind.

Bevorzugt ist erfindungsgemäß ferner ein Reaktionssystem, bei dem die ein Substrat einer Dehydrogenase darstellende organische Verbindung eine Carbonylverbindung ist.

Besonders bevorzugt ist ein Reaktionssystem, bei dem die Carbonylverbindung eine Aldehyd oder ein Keton ist.

Diese Ausführungsform der Erfindung erlaubt die erfindungsgemäß besonders bevorzugte Herstellung von technischen Alkoholen, die beispielsweise als Zwischenprodukte für die Herstellung von in Arzneimitteln einsetzbaren Wirkstoffen dienen können.

Besonders bevorzugt ist erfindungsgemäß, dass das Keton ein unsymmetrisch substituiertes Keton ist.

Diese Ausführungsform der Erfindung ist deshalb besonders bevorzugt, weil die bei einer entsprechenden Reduktion entstehenden Produkte ein Chiralitätszentrum aufweisen und mit hoher Enantioselektivität erhalten werden können. Im Allgemeinen werden die gewünschten chiralen sekundären Alkohole in enantiomerenreiner Form mit einem Enantiomerenüberschuß von >99% ee erhalten.

In einer anderen bevorzugten Ausführungsform des erfindungsgemäßen Reaktionssystems ist die ein Substrat einer Dehydrogenase darstellende organische Verbindung ein Alkohol. Diese Variante eignet sich bevorzugt zur Herstellung von kommerziell bedeutenden Carbonylverbindungen, beispielsweise Ketone mit Relevanz im Bereich der Aromachemikalien. Daneben kann die Oxidation auch zur Ausbildung von enantiomerenreinen, sekundären Alkoholen herangezogen werden, indem man von einem racemischen Alkohol als Substrat ausgeht und durch enantioselektive Oxidation das unerwünschte Enantiomer in die Ketonverbindung überführt. Das zurückbleibende, gewünschte Enantiomer kann dann entsprechend isoliert werden.

Vorzugsweise ist der Alkohol ein primärer Alkohol oder ein chiraler sekundärer Alkohol. In ersterem Fall entsteht ein Aldehyd als Produkt, wohingegen im zweiten Fall die entsprechenden Ketone gebildet werden.

Erfindungsgemäß ist auch bevorzugt, dass der Cofaktor im erfindungsgemäßen Reaktionssystem NADH, NADPH, NAD^+ oder NADP^+ ist.

Ferner betrifft die Erfindung ein Verfahren zur Herstellung des erfindungsgemäßen Polypeptids oder eines durch das erfindungsgemäße Nukleinsäuremolekül kodierten Polypeptids, wobei man den erfindungsgemäßen Wirt züchtet und das Polypeptid isoliert.

Das Polypeptid kann beispielsweise nach konventionellen Verfahren, beispielsweise nach Aufschluß entsprechender Zellen, beispielsweise mittels einer „French-Press“, durch Ionenaustausch, Größenselektion oder Affinitätschromatographie etc.

aufgereinigt werden (Coligan *et al.*, Current Protocols in Protein Science, John Wiley & Sons, Inc.). Alternativ dazu kann das erfindungsgemäße Polypeptid, wenn verknüpft mit einem Leader-Peptid, aus den Zellen ausgeschleust und aus dem Kulturüberstand aufgereinigt werden. Diese Ausführungsform bedingt, dass das natürlicherweise nicht mit einem Leader-Peptid versehene erfindungsgemäße Polypeptid gentechnisch verändert wird. Vorteilhaft an dieser Ausführungsform ist die einfachere Aufreinigung des erfindungsgemäßen Polypeptids aus dem Kulturüberstand. Die besten Verfahrensabläufe und geeignete Leader-Peptide können vom Fachmann ohne weiteres ermittelt werden.

In einer weiteren bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens wird das Polypeptid aus einer Körperflüssigkeit oder Gewebeprobe des nicht-menschlichen transgenen Tiers isoliert. Auch in dieser Ausführungsform ist das erfindungsgemäße Polypeptid vorzugsweise mit einem Leader-Peptid versehen.

In einer weiteren bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens und insbesondere wenn das nicht-menschliche transgene Tier ein Säugetier, beispielsweise eine Kuh, eine Ziege oder ein Schaf ist, ist die Körperflüssigkeit Milch oder Serum.

In einer anderen Ausführungsform betrifft die Erfindung ein Verfahren zur Herstellung einer ein Produkt einer Dehydrogenase darstellenden organischen Verbindung, wobei man eine ein Substrat einer Dehydrogenase darstellende organische Verbindung mit dem erfindungsgemäßen Polypeptid, dem erfindungsgemäßen Wirt oder mittels des erfindungsgemäßen Reaktionssystems umsetzt.

Die verschiedenen im erfindungsgemäßen Verfahren einzusetzenden Ausführungsformen der Erfindung unterscheiden sich im Prinzip dadurch, dass zu dem Polypeptid, sofern es im zellfreien in vitro System eingesetzt wird, weitere Komponenten wie Cofaktoren etc. (vgl. supra) zugesetzt werden müssen. Bei Verwendung des erfindungsgemäßen Reaktionssystems sind die notwendigen Komponenten ggf. mit Ausnahme des Substrats, vorzugsweise und vorteilhafter Weise bereits im System enthalten, ohne dass es dazu noch eines gesonderten Zusatzes bedarf.

Bevorzugt ist erfindungsgemäß ein Verfahren, das ferner den Schritt der Isolierung des Produktes der Umsetzung umfasst. Geeignete Verfahren zur Isolierung/Aufreinigung wurden vorstehend dargestellt.

In einer besonders bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens umfasst dieses ferner die Weiterverarbeitung des Produktes in ein Arzneimittel. Eine Reihe von Beschreibungen der Nutzung enantiomerenreiner Alkohole als Intermediate zur Herstellung von Pharmawirkstoffen ist in der Literatur gegeben. Eine diesbezügliche Übersicht ist u.a. enthalten in: A. Kleemann, J. Engels, B. Kutscher, D. Reichert, *Pharmaceutical Substances: Syntheses, Patents, Applications*, 4. Ausgabe, Thieme-Verlag, Stuttgart, 2001.

In einer anderen besonders bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens umfasst dieses weiter den Schritt der Weiterverarbeitung des Produktes in ein Folgeprodukt. Dabei kann die Derivatisierung sowohl durch Modifikation der Alkoholgruppe, beispielsweise durch Veresterung und anschließende Folgereaktionen, als auch durch Modifikationen der jeweiligen Substituenten erfolgen.

Dabei ist insbesondere bevorzugt, dass das erfindungsgemäße Verfahren weiter den Schritt der Formulierung des Folgeproduktes mit einem pharmazeutisch verträglichen Träger oder Exzipienten oder Verdünnungsmittel in der Herstellung eines Arzneimittels umfasst.

Beispiele für geeignete pharmazeutisch verträgliche Träger und/oder Verdünnungsmittel sind dem Fachmann bekannt und umfassen z.B. Phosphat-gepufferte Kochsalzlösungen, Wasser, Emulsionen, wie z.B. Öl/Wasser-Emulsionen, verschiedene Arten von Netzmittel oder Detergenzien, sterile Lösungen, etc. Arzneimittel, die solche Träger umfassen, können mittels bekannter konventioneller Methoden formuliert werden. Diese Arzneimittel können einem Individuum in einer geeigneten Dosis verabreicht werden. Die Verabreichung kann oral oder parenteral erfolgen, z.B. intravenös, intraperitoneal, subcutan, intramuskulär, lokal, intranasal, intrabronchial, oral oder intradermal, oder über einen Katheter an einer Stelle in einer Arterie. Präparate für eine parenterale Verabreichung umfassen sterile wässrige oder

nicht-wäßrige Lösungen, Suspensionen und Emulsionen. Beispiele für nicht-wäßrige Lösungsmittel sind Propylenglykol, Polyethylenglykol, pflanzliche Öle wie z.B. Olivenöl, und organische Esterverbindungen wie z.B. Ethyloleat, die für Injektionen geeignet sind. Wäßrige Träger umfassen Wasser, alkoholisch-wäßrige Lösungen, Emulsionen, Suspensionen, Salzlösungen und gepufferte Medien. Parenterale Träger umfassen Natriumchlorid-Lösungen, Ringer-Dextrose, Dextrose und Natriumchlorid, Ringer-Laktat und gebundene Öle. Intravenöse Träger umfassen z.B. Flüssigkeits-, Nährstoff- und Elektrolyt-Ergänzungsmittel (wie z.B. solche, die auf Ringer-Dextrose basieren. Das Arzneimittel kann außerdem Konservierungsmittel und andere Zusätze umfassen, wie z.B. antimikrobielle Verbindungen, Antioxidantien, Komplexbildner und inerte Gase. Des weiteren können, abhängig von der beabsichtigten spezifischen Verwendung, andere Wirkstoffe wie z.B. Interleukine, Wachstumsfaktoren, Differenzierungsfaktoren, Interferone, chemotaktische Proteine oder ein unspezifisches immunmodulatorisches Agens enthalten sein.

Die Art der Dosierung wird vom behandelnden Arzt entsprechend den klinischen Faktoren bestimmt. Es ist dem Fachmann bekannt, daß die Art der Dosierung von verschiedenen Faktoren abhängig ist, wie z.B. der Körpergröße bzw. dem Gewicht, der Körperoberfläche, dem Alter, dem Geschlecht oder der allgemeinen Gesundheit des Patienten, aber auch von dem speziell zu verabreichenden Mittel, der Dauer und Art der Verabreichung, und von anderen Medikamenten, die möglicherweise parallel verabreicht werden. Eine typische Dosis kann z.B. in einem Bereich zwischen 0,001 und 1000 µg liegen, wobei Dosen unterhalb oder oberhalb dieses beispielhaften Bereiches, vor allem unter Berücksichtigung der oben erwähnten Faktoren, vorstellbar sind. Im allgemeinen sollte sich bei regelmäßiger Verabreichung der erfindungsgemäßen Zusammensetzung die Dosis in einem Bereich zwischen 1 µg- und 10 mg-Einheiten pro Tag befinden. Üblicherweise werden die Wirkstoffe in diesen Zubereitungen in einer Konzentration von größer als 10 µg/ml eines physiologischen Puffers vorliegen. Sie können aber auch in fester Form in einer Konzentration von 0,1 bis 99,5 Gew.% der Gesamtmischung vorhanden sein. Im allgemeinen hat es sich als vorteilhaft erwiesen, den oder die Wirkstoffe in Gesamtmengen von etwa 0,001 bis 100 mg/kg, bevorzugt in Gesamtmengen von etwa 0,01 bis 10 mg/kg Körpergewicht je 24 Stunden, gegebenenfalls als Dauerinfusion oder in Form von mehreren Einzelgaben, zur Erzielung des

gewünschten Ergebnisses zu verabreichen. Wird die Zusammensetzung intravenös verabreicht, sollte sich die Dosis in einem Bereich zwischen 1 µg- und 10 mg-Einheiten pro Kilogramm Körpergewicht pro Tag befinden. Das Arzneimittel kann topisch, lokal oder systemisch verabreicht werden.

Erfindungsgemäß ist schließlich insbesondere bevorzugt ein Verfahren, wobei das Produkt ein enantiomerreiner Alkohol ist.

Ferner betrifft die Erfindung einen Liganden, der das erfindungsgemäße Polypeptid spezifisch bindet, wobei der Ligand kein Substrat des Polypeptids, kein Cofaktor davon und kein davon umgesetztes Produkt ist.

Der Begriff „spezifisch bindet“ bedeutet erfindungsgemäß, dass der Ligand nicht oder im wesentlichen nicht mit anderen Polypeptiden, auch solchen mit ähnlicher Primärsequenz oder ähnlicher dreidimensionaler Struktur kreuzreagiert. Kreuzreaktivität kann mit im Stand der Technik bekannten Verfahren ermittelt werden (vgl. Harlow und Lane „Antibodies, A Laboratory Manual“, CSH Press, Cold Spring Harbor, 1988). Hierzu können beispielsweise kompetitive Assays, z.B. turbidimetrische Tests eingesetzt werden, in denen der Ligand zusammen mit markiertem erfindungsgemäßen Polypeptid und einem damit kompetitierenden Polypeptid inkubiert wird, wobei letzteres in unterschiedlichen Konzentrationen eingesetzt werden kann.

In einer bevorzugten Ausführungsform ist der erfindungsgemäße Ligand ein Antikörper oder ein Fragment oder Derivat davon, ein Aptamer, oder eine niedermolekulare Substanz.

Antikörperfragmente umfassen Fv-, Fab- und F(ab')₂-Fragmente. Zu den Derivaten gehören scFvs (Harlow und Lane, loc. cit.). Antikörper können polyklonalen oder monoklonalen Ursprungs sein.

In einer besonders bevorzugten Ausführungsform des erfindungsgemäßen Rezeptors ist dieser ein monoklonaler Antikörper.

„Niedermolekulare Substanzen“ sind erfindungsgemäß natürlich vorkommende oder künstlich hergestellte Moleküle mit einem molekularen Gewicht von etwa 250 bis 1000 Da, vorzugsweise 300 bis 750 Da, besonders bevorzugt 400 bis 600 Da, oder aus Naturstoffen stammende modifizierte Moleküle mit diesem Molekulargewicht.

Die beanspruchte Erfindung umfasst weiterhin einen Primer mit einer der in Tabelle 1 dargestellten Sequenz.

Darüber hinaus betrifft die Erfindung ein Primerpaar mit in Tabelle 1 dargestellten Sequenzen, wobei der erste Primer des Primerpaares als Vorwärts-Primer und der zweite Primer des Primerpaares als Rückwärts-Primer zur Amplifizierung einer DNA-Sequenz dient.

Der erfindungsgemäße Primer (im Zusammenspiel mit einem geeigneten weiteren Primer, vorzugsweise einem weiteren in Tabelle 1 aufgeführten weiteren Primer) und das erfindungsgemäße Primerpaar können zur Amplifizierung der erfindungsgemäßen Sequenzen eingesetzt werden, vorzugsweise durch PCR oder LCR. Die Primer bzw. Primerpaare wurden aus einer Vielzahl potentiell möglicher Primer unter großem Aufwand ausgesucht. Sie erlauben neben der Amplifikation der erfindungsgemäßen Nukleinsäuresequenzen auch die Amplifikation von Sequenzen, die Enzyme aus dem Stand der Technik kodieren und sind damit vielseitig einsetzbar.

Ein anderer Gegenstand der Erfindung ist ein Kit enthaltend

- (a) das erfindungsgemäße Polypeptid;
- (b) das erfindungsgemäße Nukleinsäuremolekül;
- (c) den erfindungsgemäßen Vektor;
- (d) den erfindungsgemäßen Wirt;
- (e) den erfindungsgemäßen Liganden;
- (f) das erfindungsgemäße Reaktionssystem;
- (g) mindestens einen erfindungsgemäßen Primer; und/oder
- (h) mindestens ein erfindungsgemäßes Primerpaar.

Die Komponenten des erfindungsgemäßen Kits können einzeln oder z.T. zusammen in geeigneten Gefäßen verpackt werden. Die Komponenten können im

erfindungsgemäßen Kit beispielsweise in gefriergetrockneter Form vorliegen oder z.B. in Lösung, wobei geeignete Lösungsmittel insbesondere wässrige Lösungsmittel wie gepufferte Lösungen, z.B. phosphatgepufferte Lösungen einschließen.

Die erfindungsgemäßen Kits können vielseitig eingesetzt werden. Beispielsweise können sie zur Identifikation weiterer Alkoholdehydrogenasen oder diese kodierender Nukleinsäuren dienen, wobei vorzugsweise die erfindungsgemäßen Primer eingesetzt werden. In anderen Ausführungsformen können die erfindungsgemäßen Kits zur industriellen Herstellung des erfindungsgemäßen Enzyms oder der durch das Enzym umgesetzten Produkte eingesetzt werden. In diesen Ausführungsformen würde bevorzugt der erfindungsgemäße Wirt oder das erfindungsgemäße Reaktionssystem zum Einsatz kommen.

Die Figuren zeigen:

Fig. 1: Stand der Technik via Racematspaltung: mindestens 4-4 Schritte

Fig. 2: Übersicht des Clusters 2 (=Primergruppe 2), basierend auf 33 Sequenzen

Fig. 3: PCR-Typisierung mit Primergruppe 2 unter Verwendung verschiedener Pools

Die Beispiele erläutern die Erfindung.

Beispiel 1: Clusterung der ADHs und Primerdesign

Aus einer umfangreichen proprietären Stammsammlung wurden Stämme priorisiert, welche nach erfolgter Viabilitäts- und Reinheitskontrolle in Flüssigkultur angezogen wurden. Geerntete Zellpellets dienten als Ausgangsmaterial für das genetische Screening. Es wurden Primer für das genetische Screening nach Alkoholdehydrogenase-Genen konstruiert und anschließend auf Basis präparierter genomischer DNA ausgewählter mikrobieller Isolate mit Hilfe der PCR ausgetestet. Die NAD- oder NADP-abhängigen Alkoholdehydrogenasen werden in langkettige, mittellange und kurzkettige ADHs eingeteilt. Da die Sequenzheterogenität innerhalb dieser Gruppen signifikant ist, wurden diese basierend auf Sequenzanalysen in Clustern gruppiert. Die langkettigen ADHs wurden in drei, die mittellangen in 4 und die kurzkettigen in drei Cluster unterteilt. Anschließend wurden jeweils vier degenerierte Primer-Sets pro Cluster konstruiert, welche sich in der Nutzung

spezifischer Codons- (Codon-usage) unterscheiden, aber gegen diesselben Sequenzmotive der Cluster gerichtet sind.

Beispiel 2: Genetisches Screening nach langkettigen („long-chain“)-Alkoholdehydrogenasen

Die „long“-chain ADHs wurden auf Basis von Sequenzanalysen in drei Cluster unterteilt und die Primer unter analoger Vorgehensweise konstruiert und getestet. Allerdings wurden trotz gegenteiliger Erwartungen keine PCR-Tags amplifiziert, welche dieser Gruppe zugeordnet werden konnten.

Beispiel 3: Genetisches Screening nach mittellangen („medium-chain“)-Alkoholdehydrogenasen

Das Design von Primern, welche gegen „medium“ chain-ADHs gerichtet sind, wurde wie folgt ausgeführt: Basierend auf Sequenzanalysen wurden die „medium“ chain-ADHs in vier Cluster unterteilt. Daraufhin erfolgte die Konstruktion von jeweils vier degenerierten Primer-Sets, welche sich durch die Auswahl der „Codon-usage“ unterscheiden. Beispielhaft sei dies anhand der Organismengruppe zur Festlegung der Primergruppe 2 in Figur 2 graphisch verdeutlicht. Die Primersets wurden ausgewählt anhand von konservierten Regionen bei 33 unterschiedlichen Alkoholdehydrogenase-Sequenzen.

Diese Primergruppen, z.B. die Primergruppe 2, wurden im Anschluss zur Untersuchung diverser „Pools“ (mit genomischer DNA aus Mikroorganismen) eingesetzt. Durch Einsatz dieses Primersets konnten neuartige „medium“-chain ADH-Teilsequenzen amplifiziert, kloniert und sequenziert werden. Das zugehörige Ergebnis dieser PCR-Typisierung ist im Folgenden in Figur 3 aufgezeigt. Wie unter anderem Spur 1, 2 und 10 in Figur 3 dokumentieren, wurden hier jeweils Gensequenzen gefunden, die aufgrund der übereinstimmenden Gensequenz mit bekannten Genen von ADH-Enzymen auf eine Alkoholdehydrogenase-Aktivität hindeuten. Insgesamt wurden weitere Gensequenzen mit potentieller Alkoholdehydrogenase-Aktivität identifiziert. Die Identität der gefundenen Sequenz-Tags mit bereits bekannten ADHs lag zwischen 51- 99%.

Beispiel 4: Genetisches Screening nach speziellen „medium-chain“-Alkoholdehydrogenasen, mit Analogie zu Alkoholdehydrogenasen aus *Rhodococcus*-Stämmen

Aufgrund der interessanten Eigenschaften der bekannten, ebenfalls zu den „medium chain“-Alkoholdehydrogenasen gehörenden (S)-Alkoholdehydrogenase aus *Rhodococcus erythropolis* (S-Re-ADH; dieses Enzym zeichnet sich durch die stereoselektive Umsetzung eines breiten Spektrums an Ketonen und Ketoestern zu den korrespondierenden Hydroxyverbindungen aus) sowie weiterer aus *Rhodococcus*-Stämmen gewonnenen Alkoholdehydrogenasen wurde der Frage nachgegangen, ob es möglich ist, mit Hilfe eines genetischen Screenings neuartige ADH-Sequenzen zu identifizieren, welche eine relativ hohe Ähnlichkeit zu dieser Sequenz zeigen. Die Antriebskraft ist dabei die Annahme, daß neuartige ADHs, deren Sequenzen eine hohe Identität mit der S-Re-ADH-Sequenz teilen, ebenfalls interessante Eigenschaften besitzen könnten. Beispielsweise könnten solche neuartigen ADHs zum einen die bewährten Eigenschaften der S-Re-ADH besitzen, zum anderen aber beispielsweise über ein verändertes Substratspektrum bzw. eine erhöhte Expressionsleistung verfügen. Für die Beantwortung der obigen Frage wurden zunächst vergleichende Sequenzanalysen mit der Aminosäuresequenz dieser S-Re-ADH durchgeführt. Diese Analysen ergaben, daß es sich im Fall der S-Re-ADH um einen Vertreter des Clusters 1 der mittellangen ADHs handelt. Des Weiteren konnte innerhalb dieses Clusters eine Gruppe bestehend aus 5 Proteinsequenzen ausgemacht werden, welche die S-Re-ADH einschließt. Ausgehend von diesen 5 Sequenzen wurden degenerierte Primer unter Berücksichtigung der „codon-usage“, entsprechend dem obig beschriebenen Vorgehen, konstruiert und ausgetestet.

Um die Anzahl der durchzuführenden PCRs zu reduzieren, wurden Pools bestehend aus 24 bakteriellen Isolaten angelegt und DNA isoliert. Diese DNA wurde als Template eingesetzt. Es wurden zahlreiche Sequenz-Tags amplifiziert und sequenziert. Die Analyse der in Aminosäuresequenzen übersetzten Sequenz-Tags ergab Übereinstimmungen zur S-Re-ADH-Sequenz von etwa 84% bis 99%. Es wurden zwei Volllängengene isoliert, welche durch ein Sequenz-Tag repräsentiert werden und auf Aminosäuresequenz-Ebene eine Identität von 98% zur S-Re-ADH

zeigen. Die neuen ADHs stammen aus dem Organismus *Arthrobacter paraffineus* ATCC21317. Auf DNA-Ebene liegt die Ähnlichkeit bei 94%. Die Isolierung dieser Vollängengene wurde mit Hilfe eines sequenzhomologen Ansatzes durchgeführt.

Beispiel 5: Genetisches Screening nach kurzkettigen („short-chain“)-Alkoholdehydrogenasen

Außerdem wurden 12 Primer-Sets für die kurzkettigen ADHs, welche gegen die drei Cluster dieser Gruppe gerichtet sind, abschließend ausgetestet. Als Template wurde DNA eingesetzt, welche aus 5 Isolaten isoliert worden war, die im Aktivitäts-Screening dank ihrer ADH-Aktivität entweder 4-Chloracetophenon oder 2-Heptanon reduziert hatten. Die Identität der Aminosäure-Sequenz-Tags zu bekannten kurzkettigen ADH-Sequenzen liegt zwischen 47-84%, wobei der überwiegende Teil dieser Sequenzen eine Identität von weniger als 67% zu veröffentlichten Sequenzen zeigt.

Tabelle1: Primersequenzen, die für das Screening der die erfindungsgemäßen Alkoholdehydrogenasen kodierenden DNA-Sequenzen eingesetzt wurden

Name	Sequenz 5' → 3'	Direction	Block
ADHM1:	AAAGCATGCGGCGTTTGTCAYACNGA	Forward	A
ADHM2:	CCAATGTTTCATCGCTTGATATGBNGTRATNCC	Reverse	C
ADHM3:	TGCGGCGTCTGCCAYACBGA	Forward	A
ADHM4:	GCTTCAGGGCGTGGTAGGBVGTVAAYRCC	Reverse	C
ADHM5:	GCGGCGTCTGCCACWCSGA	Forward	A
ADHM6:	GCTTCAGGGCGTGGTAGGBSGTSAYSCC	Reverse	C
ADHM7:	AGCCTGCGGCGTCTGYCAYWCBGA	Forward	A
ADHM8:	GCTTCAGCGCTGGTAGGBSGTSAYNCC	Reverse	C
ADHM9:	GCAGCTTGCGGCATGTGYCAYACNGA	Forward	A
ADHM10:	GCCCAAGCCGGTCGTAAYNCCRCANCC	Reverse	C
ADHM11:	GGCCTGCGGCATGTGYCAYACBGA	Forward	A
ADHM12:	CCCAAGCCGGTCGTGAYRMMRCVCC	Reverse	C
ADHM13:	CCGGCATGTGCCACACSGA	Forward	A
ADHM14:	TGGCGGCCAGGCCSAYSSCSCC	Reverse	C
ADHM15:	GGCCTCCGGCATGTGYCAYACSGA	Forward	A
ADHM16:	TGGCGGCCAGGCCSAYNSCNCC	Reverse	C
ADHM17:	TTAAATGGTGCGGCATTTGYGGNWCNGA	Forward	A
ADHM18:	CAACTTAACAGCCAACATGCCDATNGKNCC	Reverse	D
ADHM19:	CAAGGTCAAGTGGTGCGGBATYTGYYG	Forward	A
ADHM20:	TGACGGCCAACATGCCRATNGKVCC	Reverse	D
ADHM21:	TGCGGCATCTGCGGSWCSGA	Forward	A
ADHM22:	CGAAGTTGACGGCGAAGAKSCCGATSGKSC	Reverse	D
ADHM23:	CAAGGTCAAGTGGTGCGGNATCTGYGG	Forward	A
ADHM24:	CGGCGAAGATGCCGATSGKNCC	Reverse	D
ADHM25:	GATTGTTAGAGTTACAGCTACAGCTATTTGYGGNWSNGA	Forward	A
ADHM26:	TGAACGGCAAACAGGCCNAYNGGNCC	Reverse	D
ADHM27:	CGCCACCGCCATCTGYGGBWSBGA	Forward	A
ADHM28:	GACGGCGAACAGGCCNAYNGGVCC	Reverse	D
ADHM29:	CACCGCCATCTGCGGSWSSGA	Forward	A
ADHM30:	GGAGTGGACGGCGAACAACSCCSAYSGGSC	Reverse	D
ADHM31:	CGCCACCGCCATCTGYGGNWSBGA	Forward	A
ADHM32:	GACGGCGAACAGGCCSAYSGGNCC	Reverse	D
ADHM39:	AGAAGAACTGGGCATTATGCCNCCNNGNYT	Forward	A
ADHM40:	TGTATCAATTGTCGGTTGATAGCCNACRAARTCNA	Reverse	D

ADHM41	ACAACGTGGTCTGTACGGNCCNTGGGG	Forward	
ADHM42	GATGGTGGGCTGGTAGCCNACRAARTCNA	Reverse	
ADHM43	GACAACGTCTGTCTACGGNCCNTGGGG	Forward	
ADHM44	AGCGCTTGATGGCGTGRTANGGNGT	Reverse	
ADHM45	GACAACGTCTGTCTACGGNCCNTGGGG	Forward	
ADHM46	GATGGTCTGGCTGGTAGCCNACRAARTCNA	Reverse	
ADHS1:	TTTGGCAGAGTTGATGTTGTTGYAAYAAYGCNG	Forward	B
ADHS2:	CCATTCAACCGCAGATTTTGWYGCRCMATA	Reverse	C
ADHS3:	CGTGTCTGACGTCGTGGTBVAYAAAYGCBG	Forward	B
ADHS4:	CACGGCGGGCTTGGWBGCGCMRTA	Reverse	C
ADHS5a	CACCGGCGGCCACCNSSGGYATSG	Forward	A
ADHS5b	CGTCTGACGTCGTCTGCVACAACGCSG	Forward	B
ADHS6	GACGGCGGGCTTGGWSGCGCMGTA	Reverse	C
ADHS7a	CGTCACCGGCGCCDCCNSSGG	Forward	A
ADHS7b	GCGTCGACGTCGTCTGCVACAACGCGG	Forward	B
ADHS8	GACGGCGGGCTTGGWSGCGCMGTA	Reverse	C
ADHS9	AAAATTGTTATTGTTACAGGAGGATCCMRNGGNATYGG	Forward	A
ADHS10	AAGCGCACATTAGCCTGCRTRTTNAYNA	Reverse	B
ADHS11	CACCGGCGGGCTGCMRNGGYATYGG	Forward	A
ADHS12	GCTGGACGTGATGTTGAYRATRBKRCC	Reverse	C
ADHS13	TCACCGGCGGGCTCCMRSGGNATCG	Forward	A
ADHS14a	CCACTAGCCCGCGTTGTTSAISA	Reverse	B
ADHS14b	GCTGCTAGGACGTGATGTTGAYGATSBKDC	Reverse	C
ADHS15	GTCATCGTCACCGGCGSSWSCMRSGG	Forward	A
ADHS16	CTAGCCGGCGTTGTTGAYSA	Reverse	B
ADHS17	TGGTTACAGGCGCAGCCMGNGGHATYGG	Forward	A
ADHS18	AATGCTTGTGTCATCATTTCCAYNSCNCCTTT	Reverse	C
ADHS19	GACCGGCGCCGGYCGYGGYATYG	Forward	A
ADHS20	AGGCCTGGGTCATCATCTCSAYRSCVCCYT	Reverse	C
ADHS21	CCGGCGCCTCGCSRVSAGVGG	Forward	A
ADHS22	CGATGGAGCCCGGGSMSACGSMGTT	Reverse	C
ADHS23	CGGCGCCTCGCCRVSGASGG	Forward	A
ADHS24	GCCCCGCGAGCTAGCCSGCGTTGTTGA	Reverse	B

SEQUENCE LISTING

<110> Degussa AG
 <120> Neue Alkoholdehydrogenasen
 <130> S-MS-IPM-PAT/Dr. Re-kö - K1419 EP
 <160> 68
 <170> PatentIn version 3.1
 <210> 1
 <211> 162
 <212> PRT
 <213> unknown
 <220>
 <221> source
 <223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
 ZF0050197= Pseudomonas oleovorans; ZF0050294= Rhodococcus;
 ZF0050330= Bacillus; ZF0051303= Bakterium; ZF0051337= Methylomonas;
 ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
 ZF0050544= Phyllobacterium rubiacearum; ZF0002852= Rhodococcus;
 ZF0050310= Arthrobacter paraffineus; ZF0002862= Streptomyces
 clavuligerus; ZF0050292= Bakterium; ZF0002031= Streptomyces;
 ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
 ZF0002437= Streptomyces; ZF0003712= Micromonospora; ZF0003765=
 Streptomyces; ZF0051305= Bakterium; ZF0003513= Actinomyces;
 ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
 ZF0002332= Streptomyces diastatochromogenes; ZF0003768= Actinomyces;
 ZF0002379= Streptomyces coelestis; ZF0002351= Nonomuraea
 roseoviolacea; ZF0003769= Actinomyces;
 <400> 1
 Gly Pro Trp Gly Cys Gly Asn Cys Trp His Cys Ser Gln Gly Leu Glu
 1 5 10 15
 Asn Tyr Cys Ser Arg Ala Gln Glu Leu Gly Ile Asn Pro Pro Gly Leu
 20 25 30
 Gly Ala Pro Gly Ala Leu Ala Glu Phe Met Ile Val Asp Ser Pro Arg
 35 40 45
 His Leu Val Pro Ile Gly Asp Leu Asp Pro Val Lys Thr Val Pro Leu
 50 55 60
 Thr Asp Ala Gly Leu Thr Pro Tyr His Ala Ile Lys Arg Ser Leu Pro
 65 70 75 80
 Lys Leu Arg Gly Gly Ser Tyr Ala Val Val Ile Gly Thr Gly Gly Leu
 85 90 95

Gly His Val Ala Ile Gln Leu Leu Arg His Leu Ser Ala Ser Thr Val
100 105 110

Ile Ala Leu Asp Val Ser Ala Asp Lys Leu Glu Leu Ala Thr Lys Val
115 120 125

Gly Ala His Glu Val Val Leu Ser Asp Lys Asp Ala Ala Glu Asn Val
130 135 140

Arg Lys Ile Thr Gly Ser Gln Gly Ala Ala Leu Val Leu Asp Phe Val
145 150 155 160

Gly Tyr

 $\langle 210 \rangle \quad 2$

<211> 128

<212> PRT

<213> unknown

$\langle 220 \rangle$

```
<221>.  source
```

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
ZF0050197= Pseudomonas oleovorans; ZF0050294= Rhodococcus;
ZF0050330= Bacillus; ZF0051303= Bakterium; ZF0051337= Methylomonas;
ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
ZF0050544= Phyllobacterium rubiacearum; ZF0002852= Rhodococcus;
ZF0050310= Arthrobacter paraffineus; ZF0002862= Streptomyces
clavuligerus; ZF0050292= Bakterium; ZF0002031= Streptomyces;
ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
ZF0002437= Streptomyces; ZF0003712= Micromonospora; ZF0003765=
Streptomyces; ZF0051305= Bakterium; ZF0003513= Actinomyces;
ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
ZF0002332= Streptomyces diastatochromogenes; ZF0003768= Actinomyces;
ZF0002379= Streptomyces coelestis; ZF0002351= Nonomuraea
roseoviolacea; ZF0003769= Actinomyces;

<400> 2

Gly Pro Trp Gly Cys Gly Asn Cys Trp His Cys Ser Gln Gly Leu Glu
1 5 10 15

Asn Tyr Cys Ser Arg Ala Gln Glu Leu Gly Ile Asn Pro Pro Gly Leu
20 25 30

Gly Ala Pro Gly Ala Leu Ala Glu Phe Met Ile Val Asp Ser Pro Arg
35 40 45

His Leu Val Pro Ile Gly Asp Leu Asp Pro Val Lys Thr Val Pro Leu
50 55 60

Thr Asp Ala Gly Leu Thr Pro Tyr His Ala Ile Lys Arg Ser Leu Pro
65 70 75 80

Lys Leu Arg Gly Gly Ser Tyr Ala Val Val Ile Gly Thr Gly Gly Leu
85 90 95

Gly His Val Thr Ile Gln Leu Leu Arg His Leu Ser Ala Ala Thr Val
100 105 110

Ile Ala Leu Asp Val Ser Ala Asp Lys Leu Glu Leu Ala Thr Lys Val
115 120 125

<210> 3

<211> 162

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050286= Corynebacterium hoagii

<400> 3

Gly Pro Trp Gly Cys Gly Arg Cys Trp His Cys Ala Gln Gly Leu Glu
1 5 10 15

Asn Tyr Cys Ser Arg Ala Arg Glu Leu Gly Ile Ala Pro Pro Gly Leu
20 25 30

Gly Ala Pro Gly Ala Ile Ala Glu Tyr Met Ile Val Asp Ser Pro Arg
35 40 45

His Leu Val Pro Ile Gly Asp Leu Asp Pro Val Thr Thr Val Pro Leu
50 55 60

Thr Asp Ala Gly Leu Thr Pro Tyr His Ala Ile Lys Arg Ser Leu Gly
65 70 75 80

Lys Leu Arg Ala Gly Ser Tyr Ala Val Val Ile Gly Thr Gly Gly Leu
85 90 95

Gly His Val Gly Ile Gln Leu Leu Arg His Leu Ser Pro Ala Arg Ile
100 105 110

Ile Ala Leu Asp Val Asn Asp Glu Lys Leu Ala Phe Ala Arg Glu Val
 115 120 125

Gly Ala His Glu Thr Val Leu Ser Asn Ala Asp Ala Ala Ala Asn Val
 130 135 140

Arg Lys Ile Thr Gly Ser Ala Gly Ala Ala Leu Val Leu Asp Phe Val
 145 150 155 160

Gly Tyr

<210> 4

<211> 161

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 4

Gly Pro Trp Gly Cys Gly Ser Cys Trp His Cys Ser Gln Gly Leu Glu
 1 5 10 15

Asn Tyr Cys Ser Arg Ala Lys Glu Leu Gly Ile Asn Pro Pro Gly Leu
 20 25 30

Gly Ala Pro Gly Ala Leu Ala Glu Phe Met Ile Val Asp Ser Pro Arg
 35 40 45

His Leu Val Pro Ile Gly Asp Leu Asp Pro Val Lys Thr Val Pro Leu
 50 55 60

Thr Asp Ala Gly Leu Thr Pro Tyr His Ala Ile Lys Arg Ser Leu Pro
 65 70 75 80

Lys Leu Arg Gly Gly Ala Tyr Ala Val Val Ile Gly Thr Gly Gly Leu
 85 90 95

Gly His Val Ala Ile Gln Leu Leu Arg His Leu Ser Ala Ala Thr Val
 100 105 110

Ile Ala Leu Asp Val Ser Ala Asp Lys Leu Val Leu Ala Thr Lys Val
 115 120 125

Gly Ala His Glu Val Val Leu Ser Asp Lys Asp Ala Ala Glu Asn Val
 130 135 140

Arg Arg Ile Thr Gly Ser Gln Gly Ala Ala Leu Val Leu Asp Phe Val
 145 150 155 160

Gly

<210> 5

<211> 70

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0004210= Actinomyces; ZF0004212= Actinomyces; ZF0004211=
 Actinomyces; ZF0003860= Actinomyces; ZF0004218= Actinomyces;
 ZF0003868= Actinomadura; ZF0004213= Actinomyces; ZF0003876=
 Actinomyces; ZF0003866= Actinomyces; ZF0003864= Actinomyces;
 ZF0003862= Actinomadura; ZF0003869= Actinomyces; ZF0003867=
 Actinomadura; ZF0004216= Actinomyces; ZF0004235= Actinomyces;
 ZF0004209= Actinomadura; ZF0004214= Actinomyces; ZF0003871=
 Actinomyces; ZF0004063= Actinomadura; ZF0004052= Actinomadura;
 ZF0006405= Streptomyces; ZF0003865= Actinomadura; ZF0004047=
 Actinomadura; ZF0004070= Actinomyces; ZF0004085= Actinomyces;
 ZF0004217= Actinomyces; ZF0004089= Actinomadura; ZF0004090=
 Actinomadura; ZF0006138= Streptomyces; ZF0004236= Actinomadura;
 ZF0051203= Bakterium;

<400> 5

Gly Pro Trp Gly Cys Gly Thr Cys Val Lys Cys Ala Glu Gly Lys Glu
 1 5 10 15

Asn Tyr Cys Leu Arg Ala Lys Glu Leu Gly Ile Ala Pro Pro Gly Leu
 20 25 30

Gly Ser Pro Gly Ala Met Ala Glu Tyr Met Ile Val Asp Asp Pro Arg
 35 40 45

His Leu Val Pro Leu Gly Gly Leu Asp Pro Val Gln Ala Val Pro Leu
 50 55 60

Thr Asp Ala Gly Leu Thr
 65 70

<210> 6

<211> 94

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
 ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
 ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
 ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
 ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
 ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
 ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
 ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
 ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821= Actinomyces;
 ZF0002359= Actinoplanes ianthinogenes; ZF0002396= Actinoplanes;
 ZF0003781= Actinomyces; ZF0003512= Actinomyces;
 ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087= Streptomyces;
 ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086= Streptomyces;
 ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
 ZF0003535= Actinomyces;

<400> 6

Cys His Thr Asp His His Ile Val Thr Gly Ala Thr Pro Met Pro Ser
 1 5 10 15

Phe Pro Val Met Gly Gly His Glu Gly Ser Gly Val Ile Thr Lys Leu
 20 25 30

Gly Pro Glu Val Lys Gly Leu Glu Val Gly Asp His Val Val Leu Ser
 35 40 45

Phe Ile Pro Ala Cys Gly Thr Cys Pro Ala Cys Ser Ala Gly His Gln
 50 55 60

Asn Leu Cys Asp Leu Gly Met Gly Leu Leu Ser Gly Gln Ala Ile Ser
 65 70 75 80

Asp Gly Thr Tyr Arg Ile Gln Ala Arg Gly Glu Asn Val Ile
 85 90

<210> 7

<211> 92

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
 ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
 ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;

ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
 ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
 ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
 ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
 ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
 ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821=
 Actinomyces; ZF0002359= Actinoplanes ianthinogenes; ZF0002396=
 Actinoplanes; ZF0003781= Actinomyces; ZF0003512= Actinomyces;
 ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087=
 Streptomyces; ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086=
 Streptomyces; ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
 ZF0003535= Actinomyces;

<400> 7

Cys His Thr Asp Asp His Ala Val Thr Gly Asp Leu Ala Val Pro Leu
 1 5 10 15

Pro Val Ile Gly Gly His Glu Gly Ala Gly Ile Val Glu Lys Val Gly
 20 25 30

Pro Gly Val Arg Asp Val Glu Val Gly Asp His Val Val Leu Ser Phe
 35 40 45

Ile Pro Ser Cys Gly Arg Cys Arg Trp Cys Ala Val Gly Gln Ser Asn
 50 55 60

Leu Cys Asp Leu Gly Ala Ile Leu Met Ala Gly Ala Gln Val Asp Gly
 65 70 75 80

Thr Tyr Arg Ala Thr Ala Arg Gly His Asp Val Gly
 85 90

<210> 8

<211> 92

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
 ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
 ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
 ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
 ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
 ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
 ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
 ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
 ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821=
 Actinomyces; ZF0002359= Actinoplanes ianthinogenes; ZF0002396=
 Actinoplanes; ZF0003781= Actinomyces; ZF0003512= Actinomyces;
 ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087=
 Streptomyces; ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086=

Streptomyces; ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
ZF0003535= Actinomyces;

<400> 8

Cys His Thr Asp Asp His Ala Val Thr Gly Asp Leu Ala Val Pro Leu
1 5 10 15

Pro Val Ile Gly Gly His Glu Gly Ala Gly Ile Val Glu Lys Val Gly
20 25 30

Pro Gly Val Arg Asp Val Glu Val Gly Asp His Val Val Leu Ser Phe
35 40 45

Ile Pro Ser Cys Gly Arg Cys Arg Trp Cys Ala Val Gly Gln Ser Asn
50 55 60

Leu Cys Asp Leu Gly Ala Ile Leu Met Ala Gly Ala Gln Val Asp Gly
65 70 75 80

Thr Tyr Arg Ala Thr Ala Arg Gly His Asp Val Gly
85 90

<210> 9

<211> 92

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821=
Actinomyces; ZF0002359= Actinoplanes ianthinogenes; ZF0002396=
Actinoplanes; ZF0003781= Actinomyces; ZF0003512= Actinomyces;
ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087=
Streptomyces; ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086=
Streptomyces; ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
ZF0003535= Actinomyces;

<400> 9

Cys His Thr Asp Asp His Ala Val Thr Gly Asp Leu Ala Val Pro Leu
1 5 10 15

Pro Val Ile Gly Gly His Glu Gly Ala Gly Ile Val Glu Lys Val Gly

20

25

30

Pro Gly Val Arg Asp Val Glu Val Gly Asp His Val Val Leu Ser Phe
 35 40 45

Ile Pro Ser Cys Gly Arg Cys Arg Trp Cys Ala Val Gly Gln Ser Asn
 50 55 60

Leu Cys Asp Leu Gly Ala Ile Leu Met Ala Gly Ala Gln Val Asp Gly
 65 70 75 80

Thr Tyr Arg Ala Thr Ala Arg Gly His Asp Val Gly
 85 90

<210> 10

<211> 92

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
 ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
 ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
 ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
 ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
 ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
 ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
 ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
 ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821=
 Actinomyces; ZF0002359= Actinoplanes ianthinogenes; ZF0002396=
 Actinoplanes; ZF0003781= Actinomyces; ZF0003512= Actinomyces;
 ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087=
 Streptomyces; ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086=
 Streptomyces; ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
 ZF0003535= Actinomyces;

<400> 10

Cys His Thr Asp Asp His Ala Val Thr Gly Asp Leu Ala Val Pro Leu
 1 5 10 15

Pro Val Ile Gly Gly His Glu Gly Ala Gly Ile Val Glu Lys Val Gly
 20 25 30

Pro Gly Val Arg Asp Val Glu Val Gly Asp His Val Val Leu Ser Phe
 35 40 45

Ile Pro Ser Cys Gly Arg Cys Arg Trp Cys Ala Val Gly Gln Ser Asn
 50 55 60

Leu Cys Asp Leu Gly Ala Ile Leu Met Ala Gly Ala Arg Val Asp Gly
 65 70 75 80

Thr Tyr Arg Ala Thr Ala Arg Gly His Asp Val Gly
 85 90

<210> 11

<211> 92

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
 ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
 ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
 ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
 ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
 ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
 ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
 ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
 ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821=
 Actinomyces; ZF0002359= Actinoplanes ianthinogenes; ZF0002396=
 Actinoplanes; ZF0003781= Actinomyces; ZF0003512= Actinomyces;
 ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087=
 Streptomyces; ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086=
 Streptomyces; ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
 ZF0003535= Actinomyces;

<400> 11

Cys His Thr Asp Asp His Ala Val Thr Gly Asp Leu Ala Val Pro Leu
 1 5 10 15

Pro Val Ile Gly Gly His Glu Gly Ala Gly Ile Val Glu Lys Val Gly
 20 25 30

Pro Gly Val Arg Asp Val Glu Val Gly Asp His Val Val Leu Ser Phe
 35 40 45

Ile Pro Ser Cys Gly Arg Cys Arg Trp Cys Ala Val Gly Gln Ser Asn
 50 55 60

Leu Cys Asp Leu Gly Ala Ile Leu Met Ala Gly Ala Gln Val Asp Gly
 65 70 75 80

Thr Tyr Arg Ala Thr Ala Arg Gly His Asp Val Gly
 85 90

<210> 12

<211> 93
 <212> PRT

<213> unknown

<220>

<221> source
 <223> ZF0050310= *Arthrobacter paraffineus*

<400> 12

Cys His Thr Asp Leu Phe Thr Lys Ser Val Leu Pro Glu Arg Leu Gly
 1 5 10 15

Pro Cys Val Phe Gly His Glu Gly Ala Gly Val Val Glu Ala Val Gly
 20 25 30

Ser Ser Ile Asp Ser Ile Ala Pro Gly Asp His Val Leu Leu Ser Tyr
 35 40 45

Arg Ser Cys Gly Val Cys Arg Gln Cys Leu Ser Gly His Arg Ala Tyr
 50 55 60

Cys Glu Ser Ser His Gly Leu Asn Ser Ser Gly Ala Arg Thr Asp Gly
 65 70 75 80

Ser Thr Pro Val Arg Arg Ser Gly Thr Pro Ile Arg Ser
 85 90

<210> 13

<211> 93

<212> PRT

<213> unknown

<220>

<221> source
 <223> ZF0002333=*Rhodococcus erythropolis*

<400> 13

Cys His Thr Asp Leu Phe Thr Lys Thr Val Leu Pro Glu Lys Leu Gly
 1 5 10 15

Pro Cys Val Phe Gly His Glu Gly Ala Gly Val Val Gln Ala Val Gly
 20 25 30

Ser Ser Ile Asp Asn Ile Ala Ala Gly Asp His Val Leu Leu Ser Tyr
 35 40 45

Arg Ser Cys Gly Val Cys Arg Gln Cys Leu Ser Asp His Arg Ala Tyr
 50 55 60

Cys Glu Ser Ser His Gly Leu Asn Ser Ser Gly Ala Arg Thr Asp Gly
 65 70 75 80

Ser Thr Pro Val Arg Arg Asn Gly Thr Pro Ile Arg Ser
 85 90

<210> 14

<211> 120

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0051303= Bakterium; ZF0051337= Methylomonas; ZF0002862= Streptomyces clavuligerus; ZF0050292= Bakterium; ZF0051305= Bakterium; ZF0003513= Actinomyces; ZF0002351= Nonomuraea roseoviolacea; ZF0003769= Actinomyces; ZF0002017= Streptomyces; ZF0051306= Bakterium; ZF0002016= Streptomyces; ZF0003504= Actinomyces; ZF0006073= Streptomyces; ZF0003770= Actinomyces; ZF0002352= Actinoplanes italicus; ZF0002378= Streptomyces aureomonopodiales; ZF0006089= Streptomyces; ZF0006106= Streptomyces; ZF0051325= Bakterium; ZF0006108= Streptomyces; ZF0002440= Streptomyces; ZF0051302= Bakterium; ZF0003532= Actinomyces; ZF0003548= Nocardiaform;

<400> 14

Cys Gly Thr Asp Arg Glu Ile Ala Ser Gly Ile Tyr Gly Trp Ala Pro
 1 5 10 15

Pro Gly Arg Glu His Leu Val Leu Gly His Glu Ser Leu Gly Arg Val
 20 25 30

Arg Thr Ala Pro Asp Gly Ser Gly Phe Ala Ala Gly Asp Leu Val Val
 35 40 45

Gly Ile Val Arg Arg Pro Asp Pro Val Pro Cys Gly Ala Cys Ala His
 50 55 60

Gly Glu Phe Asp Met Cys Arg Asn Gly Glu Tyr Val Glu Arg Gly Ile
 65 70 75 80

Lys Gln Ile Asp Gly Tyr Gly Ser Thr Ser Trp Val Val Asp Ala Asp
 85 90 95

Tyr Thr Val Lys Leu Asp Pro Ala Leu Thr Glu Val Gly Val Leu Met
 100 105 110

Glu Pro Thr Thr Val Leu Gly Gln
 115 120

<210> 15

<211> 140

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0051303= Bakterium; ZF0051337= Methylomonas; ZF0002862=
 Streptomyces clavuligerus; ZF0050292= Bakterium; ZF0051305=
 Bakterium; ZF0003513= Actinomyces; ZF0002351= Nonomuraea
 roseoviolacea; ZF0003769= Actinomyces; ZF0002017= Streptomyces;
 ZF0051306= Bakterium; ZF0002016= Streptomyces; ZF0003504=
 Actinomyces; ZF0006073= Streptomyces; ZF0003770= Actinomyces;
 ZF0002352= Actinoplanes italicus; ZF0002378= Streptomyces
 aureomonopodiales; ZF0006089= Streptomyces; ZF0006106= Streptomyces;
 ZF0051325= Bakterium; ZF0006108= Streptomyces; ZF0002440=
 Streptomyces; ZF0051302= Bakterium; ZF0003532= Actinomyces;
 ZF0003548= Nocardiaform;

<400> 15

Cys Gly Thr Asp Leu His Ile Arg Ser Trp Asp Gly Trp Ala Gln Lys
 1 5 10 15

Thr Ile Ala Thr Pro Leu Thr Leu Gly His Glu Phe Val Gly Glu Val
 20 25 30

Val Glu Thr Gly Arg Asp Val Thr Asp Ile Gln Val Gly Asp Leu Val
 35 40 45

Ser Gly Glu Gly His Leu Val Cys Gly Lys Cys Arg Asn Cys Leu Ala
 50 55 60

Gly Arg Arg His Leu Cys Arg Ala Thr Val Gly Leu Gly Val Gly Arg
 65 70 75 80

Asp Gly Ala Phe Ala Glu Tyr Val Val Leu Pro Ala Ser Asn Val Trp
 85 90 95

Val His Arg Val Pro Val Asp Leu Asp Val Ala Ala Ile Phe Asp Pro
 100 105 110

Phe Gly Asn Ala Val His Thr Ala Leu Ser Phe Pro Leu Val Gly Glu

115

120

125

Asp Val Leu Val Thr Gly Ala Gly Thr Ile Gly Ile
 130 135 140

<210> 16

<211> 138

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050197= Pseudomonas oleovorans; ZF0050294= Rhodococcus;
 ZF0050330= Bacillus, ZF0002852= Rhodococcus; ZF0050310= Arthrobacter
 paraffineus; ZF0002437= Streptomyces; ZF0003712= Micromonospora;
 ZF0003765= Streptomyces; ZF0002332= Streptomyces
 diatsatochromogenes; ZF0003768= Actinomyces; ZF0002379= Streptomyces
 coelestis; ZF0002443= Streptomyces; ZF0002442= Streptomyces;
 ZF0002436= Streptomyces; ZF0050994= Bakterium; ZF0050992= Bakterium;
 ZF0050442= Bakterium; ZF0002049= Streptomyces; ZF0006069=
 Streptomyces; ZF0006075= Streptomyces; ZF0004724= Nocardiaform;
 ZF0002392= Actinoplanes nipponensis; ZF0002356= Actinoplanes
 brasiliensis; ZF0003501= Actinomyces; ZF0051322= Bakterium;
 ZF0006078= Streptomyces; ZF0006092= Streptomyces; ZF0006090=
 Streptomyces; ZF0006084= Streptomyces; ZF0006068= Streptomyces;
 ZF0050284= Rhodococcus; ZF0050028= Agrobacterium tumefaciens;
 ZF0003540= Actinomyces; ZF0003528= Actinomyces; ZF0003529=
 Actinomyces;

<400> 16

Gly Leu Thr Ile Gly His Glu Pro Val Gly Val Ile Glu Lys Leu Gly
 1 5 10 15

Ser Ala Val Thr Gly Tyr Arg Glu Gly Gln Arg Val Ile Ala Gly Ala
 20 25 30

Ile Cys Pro Asn Phe Asn Ser Tyr Ala Ala Gln Asp Gly Ala Pro Ser
 35 40 45

Gln Asp Gly Ser Tyr Leu Val Ala Ser Gly Ala Cys Gly Cys His Gly
 50 55 60

Tyr Arg Ala Thr Ala Gly Trp Arg Phe Gly Asn Ile Ile Asp Gly Ala
 65 70 75 80

Gln Ala Glu Tyr Leu Leu Val Pro Asp Ala Gln Gly Asn Leu Ala Pro
 85 90 95

Val Pro Asp Asn Leu Ser Asp Glu Gln Val Leu Met Cys Pro Asp Ile
 100 105 110

Met Ser Thr Gly Phe Lys Gly Ala Glu Asn Ala His Ile Arg Ile Gly
 115 120 125

Asp Thr Val Ala Val Phe Ala Gln Gly Pro
 130 135

<210> 17
 <211> 144
 <212> PRT
 <213> unknown
 <220>

<221> source
 <223> ZF0050197= Pseudomonas oleovorans; ZF0050294= Rhodococcus;
 ZF0050330= Bacillus, ZF0002852= Rhodococcus; ZF0050310= Arthrobacter
 paraffineus; ZF0002437= Streptomyces; ZF0003712= Micromonospora;
 ZF0003765= Streptomyces; ZF0002332= Streptomyces
 diatsatochromogenes; ZF0003768= Actinomyces; ZF0002379= Streptomyces
 coelestis; ZF0002443= Streptomyces; ZF0002442= Streptomyces;
 ZF0002436= Streptomyces; ZF0050994= Bakterium; ZF0050992= Bakterium;
 ZF0050442= Bakterium; ZF0002049= Streptomyces; ZF0006069=
 Streptomyces; ZF0006075= Streptomyces; ZF0004724= Nocardiaform;
 ZF0002392= Actinoplanes nipponensis; ZF0002356= Actinoplanes
 brasiliensis; ZF0003501= Actinomyces; ZF0051322= Bakterium;
 ZF0006078= Streptomyces; ZF0006092= Streptomyces; ZF0006090=
 Streptomyces; ZF0006084= Streptomyces; ZF0006068= Streptomyces;
 ZF0050284= Rhodococcus; ZF0050028= Agrobacterium tumefaciens;
 ZF0003540= Actinomyces; ZF0003528= Actinomyces; ZF0003529=
 Actinomyces;

<400> 17

Cys Gly Thr Asp Leu His Ile Leu Gly Gly Asp Val Pro Glu Val Thr
 5 10 15

Asp Gly Arg Ile Leu Gly His Glu Ala Val Gly Thr Val Val Glu Val
 20 25 30

Gly Asp Gly Val Gln Thr Leu Ala Pro Gly Asp Arg Val Leu Val Ser
 35 40 45

Cys Val Thr Ala Cys Gly Thr Cys Arg Phe Cys Arg Glu Ser Arg Tyr
 50 55 60

Gly Gln Cys Leu Gly Gly Gly Gly Trp Ile Leu Gly His Leu Ile Asp
 65 70 75 80

Gly Thr Gln Ala Glu Leu Val Arg Val Pro Tyr Ala Asp Asn Ser Thr
 85 90 95

His Arg Ile Pro Asp Gly Val Ser Asp Glu Gln Met Leu Met Leu Ala
 100 105 110

Asp Ile Leu Pro Thr Ser Tyr Glu Val Gly Val Leu Asn Gly Cys Leu
 115 120 125

Arg Pro Ala Asp Val Val Val Ile Ile Gly Ala Asp Asp Arg Pro Leu
 130 135 140

<210> 18

<211> 73

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 18

Val Asp Val Val Val Asp Asn Ala Gly Phe Gly Thr His Gly Ala Phe
 1 5 10 15

Val Asp Glu Asp His Glu Arg Val Thr Ser Glu Ile Gln Leu Asn Ile
 20 25 30

Ala Thr Leu Val Glu Leu Thr His Thr Phe Pro Pro Asp Leu Leu Thr
 35 40 45

Gly Arg Gly Ala Leu Val Asn Ile Ala Ser Thr Ala Ser Phe Gln Pro
 50 55 60

Thr Pro Gly Met Ala Val Tyr Cys Ala
 65 70

<210> 19

<211> 75

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 19

Val Asp Val Val Val His Asn Ala Gly Phe Gly Thr His Gly Ala Phe
1 5 10 15

Val Asp Glu Asp Leu Glu Arg Val Thr Ser Glu Ile Gln Leu Asn Ile
20 25 30

Ala Thr Leu Val Glu Leu Thr His Thr Phe Leu Pro Asp Leu Leu Thr
35 40 45

Gly Arg Gly Ala Leu Val Asn Ile Ala Ser Thr Ala Ser Phe Gln Pro
50 55 60

Thr Pro Gly Met Ala Val Tyr Cys Ala Thr Lys
65 70 75

<210> 20

<211> 79

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0003535= Actinomyces

<400> 20

Arg Val Asp Val Val Val His Asn Ala Ala Ile Thr Gln Lys Ala Thr
1 5 10 15

Phe Arg Asp Ile Thr Pro Ala Asp Phe Glu Arg Ile Leu Arg Val Asn
20 25 30

Leu Thr Gly Val Phe Asn Leu Ser Gln Ala Val Ile Pro Leu Met Ile
35 40 45

Gln Arg Gly Gly Gly Ser Ile Val Ser Ile Ser Ser Leu Ser Ala Gln
50 55 60

Asn Gly Gly Gly Ile Phe Gly Gly Ala His Tyr Cys Ala Thr Lys
65 70 75

<210> 21

<211> 76

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0003535= Actinomyces

<400> 21

Val Asp Val Val Val Asp Asn Ala Gly Leu Ala Leu Gly Thr Ala Pro
 1 5 10 15

Ala Pro Gln Val Pro Leu Lys Asp Trp Gln Thr Met Val Asn Thr Asn
 20 25 30

Ile Thr Gly Leu Leu Asn Ile Thr His His Leu Leu Pro Thr Leu Ile
 35 40 45

Asp Arg Lys Gly Ile Val Val Asn Leu Ser Ser Val Ala Ala His Tyr
 50 55 60

Pro Tyr Thr Gly Gly Asn Val Tyr Cys Ala Ser Lys
 65 70 75

<210> 22

<211> 72

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= Arthrobacter paraffineus

<400> 22

Gln Gly Ile Gly Tyr Ala Thr Ala Lys Arg Leu Ile Ser Leu Gly Ala
 1 5 10 15

Thr Val Ala Ile Gly Asp Ile Asp Glu Ala Thr Leu Ala Arg Ala Ala
 20 25 30

Lys Asp Leu Gly Ile Arg Thr Phe Gly Arg Leu Asp Val Thr Asp Pro
 35 40 45

Ala Ser Phe Phe Asp Phe Leu Asp Thr Val Glu Gly Glu Leu Gly Pro
 50 55 60

Ile Asp Val Leu Ile Asn Asn Ala
 65 70

<210> 23

<211> 75

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 23

Gln Arg Ile Gly Leu Glu Ile Ala Arg Thr Phe Ile Lys Glu Gly Ala
1 5 10 15

Thr Val Val Leu Gly Asp Ile Asn Glu Thr Val Gly Thr Ala Ala Val
20 25 30

Ala Glu Leu Gly Gly Glu Ser Val Ala Arg Phe Ala Ser Cys Asp Val
35 40 45

Arg Asp Ser Gly Gln Val Glu Ala Met Leu Asp Leu Ala Glu Ser Ala
50 55 60

Phe Gly Pro Val Asp Val Met Met Asn Asn Ala
65 70 75

<210> 24

<211> 72

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 24

Gln Gly Ile Gly Tyr Gln Thr Ala Lys Glu Leu Ile Arg Arg Gly His
1 5 10 15

Arg Val Ala Ile Gly Asp Ile Asp Glu Ala Arg Ala Lys Glu Thr Ala
20 25 30

Ala Glu Leu Gly Val Lys Val Val Thr Arg Leu Asp Val Thr Asp Pro
35 40 45

Asp Ser Phe Lys Asp Phe Leu Asp Leu Val Glu Gly Asp Leu Gly Pro
50 55 60

Leu Asp Val Leu Ile Asn Asn Ala
65 70

<210> 25

<211> 74

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 25

Gly Ile Gly Leu Glu Ile Ala Arg Thr Phe Ile Lys Glu Gly Ala Thr
1 5 10 15

Val Val Leu Gly Asp Ile Asn Glu Thr Val Gly Thr Ala Ala Val Ala
20 25 30

Glu Leu Gly Gly Glu Ser Val Ala Arg Phe Ala Ser Cys Asp Val Arg
35 40 45

Asp Ser Gly Gln Val Glu Ala Met Leu Asp Leu Ala Glu Ser Ala Phe
50 55 60

Gly Pro Val Asp Val Ile Val Asn Asn Ala
65 70

<210> 26

<211> 74

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 26

Ile Gly Leu Glu Ile Ala Arg Thr Phe Ile Lys Glu Gly Ala Thr Val
1 5 10 15

Val Leu Gly Asp Ile Asn Glu Thr Val Gly Thr Ala Ala Val Gly Glu
20 25 30

Leu Gly Gly Glu Ser Val Ala Arg Phe Ala Ser Cys Asp Val Arg Asp
35 40 45

Ser Gly Gln Val Glu Ala Met Leu Asp Leu Ala Glu Ser Ala Phe Gly
50 55 60

Pro Val Asp Val Met Val Asn Asn Ala Gly
65 70

<210> 27

<211> 62

<212> PRT

<213> unknown

<220>

<221> source

<223> .ZF0002333= *Rhodococcus erythropolis*

<400> 27

Val Pro Val Ala Val Val Asp Leu His Ile Glu Ser Ala Lys Glu Thr
1 5 10 15

Val Ala Leu Ile Glu Ser Gln Tyr Gly Thr Pro Ala Leu Ala Leu Glu
20 25 30

Ala Asp Val Arg Asp Arg Ala Ala Val Ser Ala Ala Phe Glu Ala Thr
35 40 45

Val Ala Glu Trp Gly Arg Phe Asp Tyr Leu Val Asn Asn Ala
50 55 60

<210> 28

<211> 74

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0002333= *Rhodococcus erythropolis*

<400> 28

Leu Gly Arg Glu Ile Ala Leu Lys Leu Ala Ser Glu Gly Ala Ser Val
1 5 10 15

Val Val Asn Asp Leu Asp Pro Glu Pro Ala Ala Gln Thr Glu Arg Asp
 20 25 30

Ile Lys Ala Thr Gly Gly Gln Ala Val Ser Cys Val Gly Ser Val Ala
 35 40 45

Glu Asp Gly Phe Ala Glu Arg Phe Val Asn Thr Ala Val Glu Ser Phe
 50 55 60

Gly Gly Leu Asp Val Met Val Asn Asn Ala
 65 70

<210> 29

<211> 76

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0002333= Rhodococcus erythropolis

<400> 29

Ala Gly Leu Gly Val Glu Phe Ala His Arg Phe Ala Ala Arg Gly Ala
 1 5 10 15

Asn Leu Val Leu Val Ala Arg Arg Ala Asp Arg Leu Glu Ala Leu Ala
 20 25 30

Thr Glu Leu Arg Val Ala His Gly Ile Thr Val Thr Val Leu Pro Ala
 35 40 45

Asp Leu Ala Ala Pro Gly Val Gly Ala Thr Leu His Gln Glu Leu Thr
 50 55 60

Ser Arg Gly Ile Thr Val Thr Ser Leu Ile Asn Asn
 65 70 75

<210> 30

<211> 72

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0003535= Actinomyces

<400> 30

Pro Ala Asp Gly Tyr Gln Thr Ala Lys Glu Leu Ile Arg Arg Gly His
 1 5 10 15

Arg Val Ala Ile Val Asp Ile Asp Glu Ala Arg Ala Lys Gly Ala Ala
 20 25 30

Ala Glu Leu Gly Val Lys Val Val Thr Arg Leu Asp Val Thr Glu Pro
 35 40 45

Asp Ser Phe Thr Thr Phe Leu Asp Leu Val Glu Arg Glu Leu Gly Pro
 50 55 60

Leu Asp Ile Leu Val Asn Asn Ala
 65 70

<210> 31

<211> 67

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 31

Ala Thr Asp Gly Ala Arg Val Ala Val Val Asp Leu His Ile Glu Ser
 1 5 10 15

Ala Glu Glu Thr Val Ala Leu Ile Glu Ser Gln Tyr Gly Thr Pro Ala
 20 25 30

Leu Ala Leu Glu Ala Asp Val Arg Asp Arg Ala Ala Val Ser Ala Ala
 35 40 45

Phe Glu Ala Thr Val Ala Glu Trp Gly Arg Phe Asp Tyr Leu Val Asn
 50 55 60

Asn Ala Gly
 65

<210> 32

<211> 67

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 32

Ala Ala Asp Gly Ala Arg Val Ala Val Val Asp Leu His Ile Glu Ser
1 5 10 15

Ala Lys Glu Thr Val Ala Leu Ile Glu Ser Gln Tyr Gly Thr Pro Ala
20 25 30

Leu Ala Leu Glu Ala Asp Val Arg Asp Arg Ala Ala Val Ser Ala Ala
35 40 45

Phe Glu Ala Thr Val Ala Glu Trp Gly Arg Phe Asp Tyr Leu Val Asn
50 55 60

Asn Ala Gly
65

<210> 33

<211> 348

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 33

Met Lys Ala Ile Gln Tyr Ala Arg Ile Gly Ala Glu Pro Glu Leu Thr
1 5 10 15

Glu Ile Pro Lys Pro Glu Pro Gly Pro Gly Glu Val Leu Leu Glu Val
20 25 30

Thr Ala Ala Gly Val Cys His Ser Asp Asp Phe Ile Met Ser Leu Pro
35 40 45

Glu Glu Gln Tyr Thr Tyr Gly Leu Pro Leu Thr Leu Gly His Glu Gly
50 55 60

Ala Gly Arg Val Ala Ala Val Gly Glu Gly Val Glu Gly Leu Asp Ile
65 70 75 80

Gly Thr Asn Val Val Val Tyr Gly Pro Trp Gly Cys Gly Ser Cys Trp
85 90 95

His Cys Ser Gln Gly Leu Glu Asn Tyr Cys Ser Arg Ala Lys Glu Leu
100 105 110

Gly Ile Asn Pro Pro Gly Leu Gly Ala Pro Gly Ala Leu Ala Glu Phe
115 120 125

Met Ile Val Asp Ser Pro Arg His Leu Val Pro Ile Gly Asp Leu Asp
130 135 140

Pro Val Lys Thr Val Pro Leu Thr Asp Ala Gly Leu Thr Pro Tyr His
145 150 155 160

Ala Ile Lys Arg Ser Leu Pro Lys Leu Arg Gly Gly Ala Tyr Ala Val
165 170 175

Val Ile Gly Thr Gly Gly Leu Gly His Val Ala Ile Gln Leu Leu Arg
180 185 190

His Leu Ser Ala Ala Thr Val Ile Ala Leu Asp Val Ser Ala Asp Lys
195 200 205

Leu Glu Leu Ala Thr Lys Val Gly Ala His Glu Val Val Leu Ser Asp
210 215 220

Lys Asp Ala Ala Glu Asn Val Arg Arg Ile Thr Gly Ser Gln Gly Ala
225 230 235 240

Ala Leu Val Leu Asp Phe Val Gly Tyr Gln Pro Thr Ile Asp Thr Ala
245 250 255

Met Ala Val Ala Gly Val Gly Ser Asp Val Thr Ile Val Gly Ile Gly
260 265 270

Asp Gly Gln Ala His Ala Lys Val Gly Phe Phe Gln Ser Pro Tyr Glu
275 280 285

Ala Ser Val Thr Val Pro Tyr Trp Gly Ala Arg Asn Glu Leu Ile Glu
290 295 300

Leu Ile Asp Leu Ala His Ala Gly Ile Phe Asp Ile Ala Val Glu Thr
305 310 315 320

Phe Ser Leu Asp Asn Gly Ala Glu Ala Tyr Arg Arg Leu Ala Ala Gly

325

330

335

Thr Leu Ser Gly Arg Ala Val Val Val Pro Gly Leu
 340 345

<210> 34

<211> 348

<212> PRT

<213> unknown

<220>

<221> source

<223> ZF0050310= Arthrobacter paraffineus

<400> 34

Met Lys Ala Ile Gln Tyr Thr Arg Ile Gly Ala Glu Pro Glu Leu Thr
 1 5 10 15

Glu Ile Pro Lys Pro Glu Pro Gly Pro Gly Glu Val Leu Leu Glu Val
 20 25 30

Thr Ala Ala Gly Val Cys His Ser Asp Asp Phe Ile Met Ser Leu Pro
 35 40 45

Glu Glu Gln Tyr Thr Tyr Gly Leu Pro Leu Thr Leu Gly His Glu Gly
 50 55 60

Ala Gly Arg Val Ala Ala Val Gly Glu Gly Val Glu Gly Leu Asp Ile
 65 70 75 80

Gly Thr Asn Val Val Val Tyr Gly Pro Trp Gly Cys Gly Ser Cys Trp
 85 90 95

His Cys Ser Gln Gly Leu Glu Asn Tyr Cys Ser Arg Ala Lys Glu Leu
 100 105 110

Gly Ile Asn Pro Pro Gly Leu Gly Ala Pro Gly Ala Leu Ala Glu Phe
 115 120 125

Met Ile Val Asp Ser Pro Arg His Leu Val Pro Ile Gly Asp Leu Asp
 130 135 140

Pro Val Lys Thr Val Pro Leu Thr Asp Ala Gly Leu Thr Pro Tyr His
 145 150 155 160

Ala Ile Lys Arg Ser Leu Pro Lys Leu Arg Gly Gly Ala Tyr Ala Val

165

170

175

Val Ile Gly Thr Gly Gly Leu Gly His Val Ala Ile Gln Leu Leu Arg
 180 185 190

His Leu Ser Ala Ala Thr Val Ile Ala Leu Asp Val Ser Ala Asp Lys
 195 200 205

Leu Glu Leu Ala Thr Lys Val Gly Ala His Glu Val Val Leu Ser Asp
 210 215 220

Lys Asp Ala Ala Glu Asn Val Arg Arg Ile Thr Gly Ser Gln Gly Ala
 225 230 235 240

Ala Leu Val Leu Asp Phe Val Gly Tyr Gln Pro Thr Ile Asp Thr Ala
 245 250 255

Met Ala Val Ala Gly Val Gly Ser Asp Val Thr Ile Val Gly Ile Gly
 260 265 270

Asp Gly Gln Ala His Ala Lys Val Gly Phe Phe Gln Ser Pro Tyr Glu
 275 280 285

Ala Ser Val Thr Val Pro Tyr Trp Gly Ala Arg Asn Glu Leu Ile Glu
 290 295 300

Leu Ile Asp Leu Ala His Ala Gly Ile Phe Asp Ile Ala Val Glu Thr
 305 310 315 320

Phe Ser Leu Asp Asn Gly Ala Glu Ala Tyr Arg Arg Leu Ala Ala Gly
 325 330 335

Thr Leu Ser Gly Arg Ala Val Val Val Pro Gly Leu
 340 345

<210> 35

<211> 488

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
 ZF0050197= Pseudomonas oleovorans; ZF0050294= Rhodococcus;
 ZF0050330= Bacillus; ZF0051303= Bakterium; ZF0051337= Methylomonas;
 ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
 ZF0050544= Phyllobacterium rubiacearum; ZF0002852= Rhodococcus;

ZF0050310= *Arthrobacter paraffineus*; ZF0002862= *Streptomyces clavuligerus*; ZF0050292= *Bakterium*; ZF0002031= *Streptomyces*; ZF0002349= *Streptomyces spectabilis*; ZF0002434= *Streptomyces*; ZF0002437= *Streptomyces*; ZF0003712= *Micromonospora*; ZF0003765= *Streptomyces*; ZF0051305= *Bakterium*; ZF0003513= *Actinomyces*; ZF0050993= *Kocuria*; ZF0002018= *Streptomyces*; ZF0003767= *Actinomyces*; ZF0002332= *Streptomyces diastatochromogenes*; ZF0003768= *Actinomyces*; ZF0002379= *Streptomyces coelestis*; ZF0002351= *Nonomuraea roseoviolacea*; ZF0003769= *Actinomyces*;

```
<400> 35
gggccatggg gttgtggcaa ctgttggcac tgctcacaag gactcgagaa ctattgctct 60
cgcgcccaag aactcggaat caatcctccc ggtctcggtg caccgggcgc gttggccgag 120
ttcatgatcg tcgattctcc tcgccacctt gtcccgatcg gtgacctga cccgggtcaag 180
acggtgccgc tgaccgacgc cgggtctgacg ccgtatcacg cgatcaagcg ttctctgccg 240
aaacttcgcg gaggctcgta cgcggttgctc attggtaccg gcgggctcgg ccacgtcgcc 300
attcagctcc tccgtcacct ctcggcgtca acggtcatcg ctttggacgt gagcgcggaac 360
aagctcgaac tggcaaccaa ggtaggcgct cacgaagtgg ttctgtccga caaggacgcg 420
gccgagaacg tccgcaagat cactggaagt caaggcgccg cactggttct cgacttcggt 480
ggctacca 488
```

<210> 36
 <211> 385
 <212> DNA
 <213> unknown
 <220>

<221> source
 <223> ZF0002326= *Actinoplanes missouriensis*; ZF0003505= *Streptomyces*; ZF0050197= *Pseudomonas oleovorans*; ZF0050294= *Rhodococcus*; ZF0050330= *Bacillus*; ZF0051303= *Bakterium*; ZF0051337= *Methylomonas*; ZF0051321= *Bakterium*; ZF0050782= *Lactobacillus bulgaricus*; ZF0050544= *Phyllobacterium rubiacearum*; ZF0002852= *Rhodococcus*; ZF0050310= *Arthrobacter paraffineus*; ZF0002862= *Streptomyces clavuligerus*; ZF0050292= *Bakterium*; ZF0002031= *Streptomyces*; ZF0002349= *Streptomyces spectabilis*; ZF0002434= *Streptomyces*; ZF0002437= *Streptomyces*; ZF0003712= *Micromonospora*; ZF0003765= *Streptomyces*; ZF0051305= *Bakterium*; ZF0003513= *Actinomyces*; ZF0050993= *Kocuria*; ZF0002018= *Streptomyces*; ZF0003767= *Actinomyces*; ZF0002332= *Streptomyces diastatochromogenes*; ZF0003768= *Actinomyces*; ZF0002379= *Streptomyces coelestis*; ZF0002351= *Nonomuraea roseoviolacea*; ZF0003769= *Actinomyces*;

```
<400> 36
gggccatggg gttgtggcaa ctgttggcac tgctcacaag gactcgagaa ctattgctct 60
cgcgcccaag aactcggaat caatcctccc ggtctcggtg caccgggcgc gttggccgag 120
ttcatgatcg tcgattctcc tcgccacctt gtcccgatcg gtgacctga cccgggtcaag 180
```

acggtgccgc tgaccgacgc cgggtctgacg ccgtatcacg cgatcaagcg ttctctgccg 240
 aaacttcgcg gaggctcgta cgcggttgtc attggtaccg gcgggctcgg ccacgtcacc 300
 attcagctcc tccgtcacct ctccggcggca acggtcatcg ctttggacgt gagcgcggac 360
 aagctcgaac tggcaaccaa ggtag 385

<210> 37

<211> 486

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0050286= *Corynebacterium hoagii*

<400> 37
 ggcccttggg gttgaggacg ttgctggcac tgcgcgcagg ggctcgagaa ctactgctcc 60
 cgcgcaaggg aactcggcat cgcgccaccc ggcttggggc gcgcgggcgc gatcgccgag 120
 tacatgatcg togactcgcc gcgtcacctg gtcccgatcg gtgacctcga ccccgtcacg 180
 acggtgccgc tgaccgacgc cgggctcacc ccgtaccacg cgatcaaacg gtcgctcggc 240
 aagctccgog cgggctcgta cgcagtcgtg atcggcaccg gaggcctcgg acacgtcggc 300
 atccagctgc tccgccacct gtcccctgca cgcacatcg ccctcgacgt caacgacgag 360
 aagctcgcgt tcgcccgcga ggtcggcgcg caccgagaccg tgttgtcgaa cgcgcacgcc 420
 gcgcggaacg tccggaagat caccgggttcg gccggtgccg cgctggtoct agacttcgtc 480
 ggctac 486

<210> 38

<211> 483

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 38
 ggcccatggg gctgtggcag ctgttggcac tgctcgcaag gactcgaaaa ctactgttct 60
 cgggcaaaaag aactcggcat caatcctcct ggtctcgggtg caccgcggcgc gttggccgaa 120
 ttcatgatcg tcgattcacc tcgccacctc gtcccgatcg gcgacctcga tccggtcaag 180
 acggtgccac tgaccgacgc cgggtctgact ccgtatcacg cgatcaagcg ttcactgccg 240

aaacttcgcg gtggcgcgta cgccgtcgtc atcggtaccg gcgggtctcgg ccatgtcgcc 300
 atccaactcc tccgccacct ctccggcagca accgtcatcg cactcgacgt gagcgcggaac 360
 aagctcgtagc tggcaaccaa ggtaggcgct cacgaagtgg tcctgtccga caaggacgcg 420
 gccgagaacg tccgcaggat caccggaagt cagggcgccg cactggttct tgacttcggt 480
 ggc 483

<210> 39

<211> 210

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0004210= Actinomyces; ZF0004212= Actinomyces; ZF0004211=
 Actinomyces; ZF0003860= Actinomyces; ZF0004218= Actinomyces;
 ZF0003868= Actinomadura; ZF0004213= Actinomyces; ZF0003876=
 Actinomyces; ZF0003866= Actinomyces; ZF0003864= Actinomyces;
 ZF0003862= Actinomadura; ZF0003869= Actinomyces; ZF0003867=
 Actinomadura; ZF0004216= Actinomyces; ZF0004235= Actinomyces;
 ZF0004209= Actinomadura; ZF0004214= Actinomyces; ZF0003871=
 Actinomyces; ZF0004063= Actinomadura; ZF0004052= Actinomadura;
 ZF0006405= Streptomyces; ZF0003865= Actinomadura; ZF0004047=
 Actinomadura; ZF0004070= Actinomyces; ZF0004085= Actinomyces;
 ZF0004217= Actinomyces; ZF0004089= Actinomadura; ZF0004090=
 Actinomadura; ZF0006138= Streptomyces; ZF0004236= Actinomadura;
 ZF0051203= Bakterium;

<400> 39
 ggaccgtggg gctgcgccac gtgcgtaag tgcgcgagg gcaaggagaa ctactgcctg 60

cgcgccaagg aactcgcat cgccccgcc ggactcggt cgcccgccgc catggccgag 120

tacatgatcg tcgacgacc gcgccacct gtgcccgtcg gcgggtctga ccgggtccag 180

gccgtgccgc tcaactgacgc gggcctgaca 210

<210> 40

<211> 282

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
 ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
 ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
 ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
 ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;

ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
 ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
 ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
 ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821=
 Actinomyces; ZF0002359= Actinoplanes ianthinogenes; ZF0002396=
 Actinoplanes; ZF0003781= Actinomyces; ZF0003512= Actinomyces;
 ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087=
 Streptomyces; ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086=
 Streptomyces; ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
 ZF0003535= Actinomyces;

<400> 40
 tgtcacaccg atcaccacat cgtcacccgc ggcaccccca tgccgtcggt cccggtcatg 60
 ggcggggcag aggggttcggg cgtcatcacc aagctcggcc ctgaggtcaa gggactggag 120
 gtccggcgacc acgtcggttct gtccttcatt cgggcttggtg gaacctgtcc ggcgtgttcg 180
 gccggggcatc agaattcttg tgacctcggg atgggcctcc tcagcggcca agccatcagc 240
 gacggcacgt accggatcca ggctcgcggc gaaaacgtga tc 282

<210> 41

<211> 276

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
 ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
 ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
 ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
 ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
 ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
 ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
 ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
 ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821=
 Actinomyces; ZF0002359= Actinoplanes ianthinogenes; ZF0002396=
 Actinoplanes; ZF0003781= Actinomyces; ZF0003512= Actinomyces;
 ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087=
 Streptomyces; ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086=
 Streptomyces; ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
 ZF0003535= Actinomyces;

<400> 41
 tgccataccg acgatcatgc tgtgaccggt gatctggcag tcccactccc cgtgatcggt 60
 ggccacgaag gcgcgggcat agtgagaaaa gtcggccccc gcgtgcgaga cgtcgaggta 120
 ggcgatcacg tcgtcctctc cttcattccc tcgtgtggac gctgccgttg gtgcgcagtc 180
 ggacagagca acctctgcga cctcggcgcc attctgatgg ccggcgacaca ggtcgacggg 240
 acgtaccgag cgacagctcg cgggcacgac gtccga 276

<210> 42

<211> 276

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821= Actinomyces;
ZF0002359= Actinoplanes ianthinogenes; ZF0002396= Actinoplanes;
ZF0003781= Actinomyces; ZF0003512= Actinomyces;
ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087= Streptomyces;
ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086= Streptomyces;
ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
ZF0003535= Actinomyces;

<400> 42

tgccatacag acgatcatgc tgtgaccggt gatctggcag tccactccc cgtgatcggt 60
ggccacgaag ggcggggcat agtggagaaa gtgggcccgc gcgtgcgaga cgctcgaggta 120
ggcgatcacg togtcctctc cttcattccc tcgtgtggac gctgccgttg gtgcgcagtc 180
ggacagagca acctctgcga cctcggcgcc attctgatgg ccggcgacaca ggctcgacggg 240
acgtaccgcg cgacagctcg cgggcacgac gtcgga 276

<210> 43

<211> 276

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821= Actinomyces;
ZF0002359= Actinoplanes ianthinogenes; ZF0002396= Actinoplanes;
ZF0003781= Actinomyces; ZF0003512= Actinomyces;
ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087= Streptomyces;
ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086= Streptomyces;

Streptomyces; ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
ZF0003535= Actinomyces;

<400> 43
tgtcatactg acgatcatgc tgtgaccggt gatctggcag tcccactccc cgtgatcggt 60
ggccacgaag gcgcgggcat agtggagaaa gtcggccccg gcgtgcgaga cgtcgaggta 120
ggcgatcacg tcgtcctctc cttcattccc tcgtgtggac gctgccgttg gtgcgcagtc 180
ggacagagca acctctgcga cctcggcgcc attctgatgg ccggcgcaca ggtcgacggg 240
acgtaccgcg cgacagctcg cgggcacgac gtcgga 276

<210> 44

<211> 276

<212> DNA

<213> unknown

<220>

<223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821=
Actinomyces; ZF0002359= Actinoplanes ianthinogenes; ZF0002396=
Actinoplanes; ZF0003781= Actinomyces; ZF0003512= Actinomyces;
ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087=
Streptomyces; ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086=
Streptomyces; ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
ZF0003535= Actinomyces;

<400> 44
gtcacaccg acgatcatgc tgtgaccggt gatctggcag tcccactccc cgtgatcggt 60
ggccacgaag gcgcgggcat agtggagaaa gtcggccccg gcgtgcgaga cgtcgaggta 120
ggcgatcacg tcgtcctctc cttcattccc tcgtgtggac gctgccgttg gtgcgcagtc 180
ggacagagca acctctgcga cctcggcgcc attctgatgg ccggcgcacg ggtcgacggg 240
acgtaccgcg cgacagctcg cgggcacgac gtcgga 276

<210> 45

<211> 276

<212> DNA

<213> unknown

<220>

<221> source
 <223> ZF0002326= Actinoplanes missouriensis; ZF0003505= Streptomyces;
 ZF0051321= Bakterium; ZF0050782= Lactobacillus bulgaricus;
 ZF0050544= Phyllobacterium rubiacearum; ZF0002031= Streptomyces;
 ZF0002349= Streptomyces spectabilis; ZF0002434= Streptomyces;
 ZF0050993= Kocuria; ZF0002018= Streptomyces; ZF0003767= Actinomyces;
 ZF0003764= Streptomyces; ZF0002331= Actinoplanes philippinensis;
 ZF0002441= Streptomyces; ZF0051307= Bakterium; ZF0051301= Bakterium;
 ZF0051240= Bakterium; ZF0002333= Rhodococcus erythropolis;
 ZF0003713= Micromonospora; ZF0004980= Streptomyces; ZF0004821=
 Actinomyces; ZF0002359= Actinoplanes ianthinogenes; ZF0002396=
 Actinoplanes; ZF0003781= Actinomyces; ZF0003512= Actinomyces;
 ZF0006093= Streptomyces; ZF0006103= Streptomyces; ZF0006087=
 Streptomyces; ZF0050446= Bakterium; ZF0050445= Bakterium; ZF0006086=
 Streptomyces; ZF0002322= Rhodococcus; ZF0003538= Actinomyces;
 ZF0003535= Actinomyces;

<400> 45
 tgtcacactg acgatcatgc tgtgaccggt gatctggcag tccactccc cgtgatcgggt 60
 ggccacgaag ggcggggcat agtggagaaa gtcggcccg gcgtgcgaga cgtcgaggta 120
 ggcatcacg tcgtcctctc cttcattccc tcgtgtggac gctgccgttg gtgcgcagtc 180
 ggacagagca acctctgcga cctcggcgcc attctgatgg cggcgccaca ggtcgacggg 240
 acgtaccgcg cgacagctcg cgggcacgac gtcgga 276

<210> 46
 <211> 279
 <212> DNA
 <213> unknown
 <220>

<221> source
 <223> ZF0050310= Arthrobacter paraffineus

<400> 46
 tgccacacag atctgttcac gaagtcggtg ctaccgaaa ggctcggccc ctgcgtgttc 60
 gggcacgaag gagcgggggt ggtcgaggcc gtcggctcgt cgatcgacag cattgcgccc 120
 ggtgatcacg tggtgctgag ctaccgcagt tgcggtgtgt gcaggcagtg cctcagcggg 180
 catcgggcgt actgcgaaag ctcacacggg ctcaacagct ctggcgcacg caccgacggc 240
 tcgacgccgg tccggcgaag cggaactccg atacggtcc 279

<210> 47
 <211> 279
 <212> DNA
 <213> unknown
 <220>

<221> source
 <223> ZF0002333= Rhodococcus erythropolis

<400> 47
 tgtcatactg atctgttcac gaagacgggtg ctaccggaaa agctcggggcc ctgcgtgttc 60
 ggacacgaag gcgccggcgt cgtgcaagcc gttggctcgt cgatcgacaa catcgoggct 120
 ggtgatcacg tattgctgag ctaccgcagt tgcgggtgtat gcaggcaatg tctcagcgac 180
 catcggggcgt actgcgaaag ctacacgggg ctcaacagct ctggcgcacg caccgacggc 240
 tcgacgccgg tccggcgaaa cggaactccg atacgggtcc 279

<210> 48

<211> 360

<212> DNA

<213> unknown

<220>

<221> source
 <223> ZF0051303= Bakterium; ZF0051337= Methylomonas; ZF0002862= Streptomyces clavuligerus; ZF0050292= Bakterium; ZF0051305= Bakterium; ZF0003513= Actinomyces; ZF0002351= Nonomuraea roseoviolacea; ZF0003769= Actinomyces; ZF0002017= Streptomyces; ZF0051306= Bakterium; ZF0002016= Streptomyces; ZF0003504= Actinomyces; ZF0006073= Streptomyces; ZF0003770= Actinomyces; ZF0002352= Actinoplanes italicus; ZF0002378= Streptomyces aureomonopodiales; ZF0006089= Streptomyces; ZF0006106= Streptomyces; ZF0051325= Bakterium; ZF0006108= Streptomyces; ZF0002440= Streptomyces; ZF0051302= Bakterium; ZF0003532= Actinomyces; ZF0003548= Nocardiaform;

<400> 48
 tgcgggacgg accgcgagat cgcctcgggc atctacgggt gggcgccgcc gggacgcgaa 60
 cacctcgtcc tcgggcacga atcgctgggc cgagtacgca ccgcgcccga cggcagcggt 120
 ttgcgccgcc gtgatctcgt cgtcgggatc gtgcgcaggc ccgatccggt gccgtgcggg 180
 gcgtgtgcmc acggtgagtt cgacatgtgc cgcaacgggt agtacgtcga gcgcggggatc 240
 aagcagatcg acgggtacgg gtcgacgtcg tgggtggtgg acgccgacta cacggtcaag 300
 ctggacccgg cgctcaccga ggtgggtgtg ctgatggaac cgacgacggt gcttggccaa 360

<210> 49

<211> 421

<212> DNA

<213> unknown

<220>

<221> source
 <223> ZF0051303= Bakterium; ZF0051337= Methylomonas; ZF0002862= Streptomyces clavuligerus; ZF0050292= Bakterium; ZF0051305= Bakterium; ZF0003513= Actinomyces; ZF0002351= Nonomuraea roseoviolacea; ZF0003769= Actinomyces; ZF0002017= Streptomyces; ZF0051306= Bakterium; ZF0002016= Streptomyces; ZF0003504= Actinomyces; ZF0006073= Streptomyces; ZF0003770= Actinomyces; ZF0002352= Actinoplanes italicus; ZF0002378= Streptomyces aureomonopodiales; ZF0006089= Streptomyces; ZF0006106= Streptomyces; ZF0051325= Bakterium; ZF0006108= Streptomyces; ZF0002440= Streptomyces; ZF0051302= Bakterium; ZF0003532= Actinomyces; ZF0003548= Nocardiaform;

<400> 49
 tgtggtaccg acctgcacat ccggtcctgg gacggatggg cgcagaagac catcgccacc 60
 ccgctcacgc tcggccacga gttcgtcggc gaggtcgtcg agaccggccg cgacgtgacc 120
 gacatccagg tcggcgacct ggtcagcggc gagggccacc tgggtctgagg caagtgcgcg 180
 aactgcctgg ccggccgccc tcacctgtgc cgcgcgaccg tcggcctcgg tgcggccggt 240
 gacggcgcct tcgccgagta cgtgggtgctg cccgcctcca acgtgtgggt gcaccgggtg 300
 ccggtcgacc tcgacgtcgc cgcgatcttc gaccggttcg gcaacgcggt gcacaccgcg 360
 ctctccttcc cgctcgtcgg cgaggacgtg ctggtcaccg gtgccggtac catcggcac 420
 t 421

<210> 50

<211> 414

<212> DNA

<213> unknown

<220>

<221> source
 <223> ZF0050197= Pseudomonas oleovorans; ZF0050294= Rhodococcus; ZF0050330= Bacillus; ZF0002852= Rhodococcus; ZF0050310= Arthrobacter paraffineus; ZF0002437= Streptomyces; ZF0003712= Micromonospora; ZF0003765= Streptomyces; ZF0002332= Streptomyces diatsatochromogenes; ZF0003768= Actinomyces; ZF0002379= Streptomyces coelestis; ZF0002443= Streptomyces; ZF0002442= Streptomyces; ZF0002436= Streptomyces; ZF0050994= Bakterium; ZF0050992= Bakterium; ZF0050442= Bakterium; ZF0002049= Streptomyces; ZF0006069= Streptomyces; ZF0006075= Streptomyces; ZF0004724= Nocardiaform; ZF0002392= Actinoplanes nipponensis; ZF0002356= Actinoplanes brasiliensis; ZF0003501= Actinomyces; ZF0051322= Bakterium; ZF0006078= Streptomyces; ZF0006092= Streptomyces; ZF0006090= Streptomyces; ZF0006084= Streptomyces; ZF0006068= Streptomyces; ZF0050284= Rhodococcus; ZF0050028= Agrobacterium tumefaciens; ZF0003540= Actinomyces; ZF0003528= Actinomyces; ZF0003529= Actinomyces;

<400> 50
 ggcctgacga tcggccatga accggtgggc gtcacgaaa agctgggcag cgccgtgacg 60

gggtaccgcg agggccaacg cgtgatcgcc ggcgcgatct gcccactt caattcgat 120
 gccgcgcagg atggcgcgcc gtgcaggat ggcagctacc tgggtggccag cggcgcatgc 180
 ggctgccatg gataccgggc cacggccggc tggcgctttg gcaacatcat cgatggcgcc 240
 caggccgaat acctgctggt tcccgatgcg cagggcaatc tggcgccggt tccggacaac 300
 ctgagcgatg aacagggtgct gatgtgcccg gacatcatgt ccaccggctt caaaggcgca 360
 gagaacgcac acatccgcat cggcgacacg gtggcggtat ttgcgcaggg acca 414

<210> 51

<211> 432

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0050197= Pseudomonas oleovorans; ZF0050294= Rhodococcus;
 ZF0050330= Bacillus, ZF0002852= Rhodococcus; ZF0050310= Arthrobacter
 paraffineus; ZF0002437= Streptomyces; ZF0003712= Micromonospora;
 ZF0003765= Streptomyces; ZF0002332= Streptomyces
 diatsatochromogenes; ZF0003768= Actinomyces; ZF0002379= Streptomyces
 coelestis; ZF0002443= Streptomyces; ZF0002442= Streptomyces;
 ZF0002436= Streptomyces; ZF0050994= Bakterium; ZF0050992= Bakterium;
 ZF0050442= Bakterium; ZF0002049= Streptomyces; ZF0006069=
 Streptomyces; ZF0006075= Streptomyces; ZF0004724= Nocardiaform;
 ZF0002392= Actinoplanes nipponensis; ZF0002356= Actinoplanes
 brasiliensis; ZF0003501= Actinomyces; ZF0051322= Bakterium;
 ZF0006078= Streptomyces; ZF0006092= Streptomyces; ZF0006090=
 Streptomyces; ZF0006084= Streptomyces; ZF0006068= Streptomyces;
 ZF0050284= Rhodococcus; ZF0050028= Agrobacterium tumefaciens;
 ZF0003540= Actinomyces; ZF0003528= Actinomyces; ZF0003529=
 Actinomyces;

<400> 51

tgcgggacgg acctgcacat cctcggaggt gacgtccccg aggtgaccga cgggcgaatc 60
 ctggggccacg aggcgctcgg gaccgtggtc gaggtgggcg acggcgtaca gacactcgcg 120
 ccggggcgatc gcgtgctcgt ctcgtgtgtc accgcatgcg gtacgtgcgc gttctgccgc 180
 gagagccgct acgggcaatg cctcggaggg ggcggctgga tcctcggaca cctgatcgac 240
 ggcacccagg ccgaactcgt ccgagttccg tacgcccaca attcgaccca ccgcatcccc 300
 gacggtgtga gcgacgagca gatgctcatg ctcgccgaca tcctgcccac ctcctacgag 360
 gtccgtgttc tcaacggctg tctccggccg gcggacgtcg tcgtcatcat cggggccgac 420
 gatcggcctc tt 432

<210> 52

<211> 220

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 52
 cgtcgacgtc gtcgtcgaca acgcgggatt cggaacacac ggggcattcg tggacgaaga 60
 tcacgagcgc gtcacgtccg agattcagct caacatcgcc acgctggtcg agctgacaca 120
 cacattcccg cccgaccttc tcaccggccg cggagcactg gtcaacatcg ccagcacagc 180
 gtcgttccag ccgacaccgg gcatggccgt ctactgcgct 220

<210> 53

<211> 226

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 53
 cgtcgacgtc gtcgtccaca acgccggatt cggaacacac ggggcattcg tggacgaaga 60
 tctcgagcgc gtcacgtccg agattcagct caacatcgcc acgctggtcg agctgacaca 120
 cacattcctg cccgaccttc tcaccggccg cggagcactg gtcaacatcg ccagcacagc 180
 gtcgttccag ccgacaccgg gcatggccgt ctactgogcc accaag 226

<210> 54

<211> 237

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0003535= *Actinomyces*

<400> 54
 cgtgtcgacg tcgtggtgca caatgctgcg atcactcaaa aggccacttt tcgcgacatt 60
 accccgcgcg attttgagcg catcctgcfg gtgaacctga ccggcgtctt caacctgagc 120
 caagccgtca ttcccttgat gattcagcgc ggcggaggaa gcatcgtctc gatttctcgc 180

ctgtcggcgc agaacggcgg ggggatcttc ggcggcgccc actattgcgc aaccaag

237

<210> 55

<211> 229

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0003535= Actinomyces

<400> 55

cgtcgacgtc gtcgtcgaca acgccggtct ggcactgggc acggcccccg cgcgcaggt 60

gccgctaaag gactggcaga ccatggtgaa caccaacatc accggtctac tgaacatcac 120

caccatctc ctgccgacac tgatcgaccg taaaggtatc gtcgtcaacc tttcgtctgt 180

tgccgcgcac tatccctata cgggcggcaa tgtatactgc gcctccaag 229

<210> 56

<211> 216

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0050310= Arthrobacter paraffineus

<400> 56

gaggggatcg gatacgccac cgcgaagcgg ctgatcagcc tgggtgcgac ggtcgcgac 60

ggcgacatcg acgaagccac tctcgcgcga gcagccaagg atttgggcat ccgcacgttc 120

gggcgcctcg acgtcaccga cccgcctcg ttcttcgact tctcgcacac cgtcgaaggt 180

gaactcggcc cgatcgacgt gctgatcaac aacgcg 216

<210> 57

<211> 225

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0080310= Arthrobacter paraffineus

<400> 57

cagcggatcg ggctcgaaat tgcgcgcacc ttcatcaagg aaggcgcgac cgtcgttctc 60
 ggcgacatca acgaaaccgt gggaacggct gcggtcgccg aactcgggtg agagtcggtc 120
 gcccgtttcg cttcctgcga cgtgcgtgac tccggacagg tcgaggccat gctcgatctg 180
 gccgaaagcg ctttcgggtcc agtcgatgtc atgatgaaca acgcg 225

<210> 58

<211> 216

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0080310= *Arthrobacter paraffineus*

<400> 58
 caggggatcg gctaccagac cgcgaaggag ctgatccgac gaggtcaccg cgtggccatc 60
 ggcgacatcg acgaggcacg tgctaaggag accgccgccg aactgggggt taaggttgtc 120
 accgcctcg atgtcacoga ccctgactcg ttcaaagact ttctcgacct agtcgaggga 180
 gacctcggcc ccctcgacgt gctgatcaac aacgcg 216

<210> 59

<211> 222

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0080310= *Arthrobacter paraffineus*

<400> 59
 gggatcgggc tcgaaattgc gcgcaccttc atcaaggaag gcgcgaccgt cgttctcggc 60
 gacatcaacg aaaccgtggg aacggctgcg gtcgccgaac tcggtggaga gtcggtcgcc 120
 cgtttcgctt cctgcgacgt gcgtagctcc ggacaggctg aggccatgct cgatctggcc 180
 gaaagcgctt tcggtccagt cgatgtcatc gtgaacaacg cg 222

<210> 60

<211> 222

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0080310= *Arthrobacter paraffineus*

<400> 60

atcgggctcg aaattgcgcg caccttcacg aaggaaggcg cgaccgtcgt tctcggcgac 60
 atcaacgaaa ccgtgggaac ggctgcggtc ggcgaactcg gtggagagtc ggtcgcccgt 120
 ttcgcttcct gcgacgtgcg tgactccgga caggctcgagg ccatgctcga tctggccgaa 180
 agcgctttcg gtccagtcga tgtcatggtc aacaacgccg gc 222

<210> 61

<211> 186

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0002333= *Rhodococcus erythropolis*

<400> 61

gtgccggtcg cggctcgtgga ccttcacatc gaaagtgcaa aggagaccgt cgcacttata 60
 gaatcgagc acggcacacc cgcgctcgcc cttgaggcgg atgtgcgcga ccgcgcgcgc 120
 gtgagcgccg ctttcgaagc caccgtcgcc gaatggggac gcttcgacta cctcgtcaac 180
 aacgcc 186

<210> 62

<211> 222

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0002333= *Rhodococcus erythropolis*

<400> 62

ctcggccgtg aaatcgctct caagctcgct tccgaaggcg cctcggtagt ggtcaacgac 60
 ctogatcccg aacctgcgcg tcagaccgag cgcgatatca aagccacagg tggacaggct 120
 gtctcgtgcg tcggctccgt tgccgaggac gggttcgccg aacgcttcgt gaacactgcc 180
 gtcgaatcat tcggcgggact cgacgtcatg gtgaacaacg cg 222

<210> 63

<211> 231

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0002333= *Rhodococcus erythropolis*

<400> 63
 gcggggctcg gagtggaatt cgctcacgcg ttcgcgcgtc gcggtgcaaa totggttctc 60
 gtcgccaggc gggcagatcg cctcgaagcc ctogctaccg aactccgcgt cgcccacggc 120
 atcacagtca cagttctgcc tgccgacctg gcggcgcccg gcgtcggcgc aacactgcac 180
 caggagctga caagccgtgg catcacgcgc aactcgtga tcaacaacgc c 231

<210> 64

<211> 216

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0003535= *Actinomyces*

<400> 64
 ccagcggacg gctatcagac agcgaaggag ttgattcgac gaggccaccg ggtcgocato 60
 gtcgacatcg acgaggcacg tgccaagggg gccgcgcgcg aactcggggg gaaggctcgc 120
 acccgactcg acgtcaccga acctgactcg ttcacaacgt ttctggacct ggtcgaacgt 180
 gaactcggac ccctcgacat cctggtcaac aacgcg 216

<210> 65

<211> 201

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 65
 gccacggacg gtgcccgcgt cgcggtcgtg gaccttcaca togaaagtgc agaggagacc 60
 gtcgcactta togaatcgca gtacggcaca cccgcgcgtc cccttgaggc cgatgtgcgc 120
 gaccgcgcgc ccgtgagcgc cgctttcgaa gccacgcgtc ccgaatgggg acgcttcgac 180

tacctcgtca acaacgccgg c

201

<210> 66

<211> 201

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 66

gccgcggacg gtgcccgcgt cgcggtcgtg gaccttcaca tcgaaagtgc aaaggagacc 60

gtcgcaactta tcgaatcgca gtacggcaca cccgcgctcg cccttgaggc cgatgtgcgc 120

gaccgcgccg ccgtgagcgc cgctttcgaa gccaccgtcg ccgaatgggg acgcttcgac 180

tacctcgtca acaacgccgg c 201

<210> 67

<211> 1047

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 67

atgaaggcaa tccagtacgc gagaatcggc gcagaaccgc aactcacgga gattocccaaa 60

cccagagccc gtccagggtga agtgctcctg gaagtcaccg ctgccggcgt ctgccactcg 120

gacgacttca tcatgagcct gcccgaaagag cagtacacct acggccttcc tctcacgctc 180

ggccacgaag gcgcccggcc ggtcgccgcc gtgcggcgagg gcgtcgaagg actcgacatc 240

ggaaccaatg togtcgtcta cggaccctgg ggctgtggca gctgttggca ctgctcgcaa 300

ggaactcgaaa actactgttc tcgggcaaaa gaactcggca tcaatcctcc tgggtctcgg 360

gcacccggcg cgttgggcca attcatgata gtcgattcac ctgccacct cgtcccgatc 420

ggcgacctcg atccgggtcaa gacgggtgcc ctgaccgacg ccggtctgac tccgtatcac 480

gcgatcaagc gttcactgcc gaaacttcgc ggtggcgcggt acgccgtcgt catcggtacc 540

ggcgggtctcg gccatgtcgc catccaaactc ctccgccacc tctcggcagc aaccgtcatc 600

gcaactcgacg tgagcgcgga caagctcgaa ctggcaacca aggtaggcgc tcacgaagtg 660

gtcctgtccg acaaggacgc ggccgagaac gtccgcagga tcaccggaag tcagggcgcc 720

gcactggttc tcgacttcgt cggctatcag cccaccatcg acaccgcgat ggctgtcgcc 780
 ggcgtcggat cggacgtcac gatcgtcggg atcggcgacg ggcaggccca tgccaaagtc 840
 gggttcttcc aaagtcctta cgaggcttct gtgacagttc cgtactgggg tgcccgcac 900
 gagctgatcg aattgatcga cctgggcgcac gccggcatct tcgacatcgc ggtggagacc 960
 ttcagtctcg acaacggcgc cgaagcgtat cgacgactgg ccgccggaac gctcagcggc 1020
 cgcgcgggttg tggtccttg tctgtag 1047

<210> 68

<211> 1047

<212> DNA

<213> unknown

<220>

<221> source

<223> ZF0050310= *Arthrobacter paraffineus*

<400> 68
 atgaaggcaa tccagtacac gagaatcggc gcagaaccgc aactcacgga gattcccaaa 60
 cccgagcccg gtccagggtga agtgctcctg gaagtcaccg ctgccggcgt ctgccactcg 120
 gacgaattca tcatgagcct gccgaagag cagtacacct acggccttcc tctcacgctc 180
 ggccacgaag gcgccggccg ggtcgccgcc gtcggcgagg gcgtcgaagg actcgacatc 240
 ggaaccaatg tcgtcgtcta cggaccctgg ggctgtggca gctgttggca ctgctcgcaa 300
 ggactcgaaa actactgttc tcggggcaaaa gaactcggca tcaatcctcc tggctcgggt 360
 gcacccggcg cgttggccga attcatgata gtcgattcac ctgccacct cgtcccgatc 420
 ggcgacctcg atccggtcaa gacggtgcca ctgaccgacg ccggtctgac tccgtatcac 480
 gcgatcaagc gttcactgcc gaaacttcgc ggtggcgcggt acgccgtcgt catcggtacc 540
 ggcggtctcg gccatgtcgc catccaactc ctccgccacc totcggcagc aaccgtcatc 600
 gcactcgacg tgagcgcgga caagctcgaa ctggcaacca aggtaggcgc tcacgaagtg 660
 gtcctgtccg acaaggacgc ggccgagaac gtccgcagga tcaccggaag tcagggcgcc 720
 gcactggttc tcgacttcgt cggctatcag cccaccatcg acaccgcgat ggctgtcgcc 780
 ggcgtcggat cggacgtcac gatcgtcggg atcggcgacg ggcaggccca tgccaaagtc 840
 gggttcttcc aaagtcctta cgaggcttct gtgacagttc cgtactgggg tgcccgcac 900
 gagctgatcg aattgatcga cctgggcgcac gccggcatct tcgacatcgc ggtggagacc 960
 ttcagtctcg acaacggcgc cgaagcgtat cgacgactgg ccgccggaac gctcagcggc 1020
 cgcgcgggttg tggtccttg totgtag 1047

Patentansprüche

1. Polypeptid mit der biologischen Aktivität einer NAD- oder NADP-abhängigen Alkoholdehydrogenase, das eine der nachfolgenden Aminosäuresequenzen umfasst oder aufweist: die Sequenz der SEQ ID NO.: 1, die Sequenz der SEQ ID NO.: 2, die Sequenz der SEQ ID NO.: 3 oder eine Sequenz, die zu mindestens 90% identisch mit der Sequenz von der SEQ ID NO.: 3 ist, die Sequenz der SEQ ID NO.: 4, die Sequenz der SEQ ID NO.: 5 oder eine Sequenz, die zu mindestens 90% identisch mit der Sequenz der SEQ ID NO.: 5 ist, die Sequenz der SEQ ID NO.: 6 oder eine Sequenz, die zu mindestens 90% identisch mit der Sequenz der SEQ ID NO.: 6 ist, die Sequenz der SEQ ID NO.: 7 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 7 ist, die Sequenz der SEQ ID NO.: 8 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 8 ist, die Sequenz der SEQ ID NO.: 9 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 9 ist, die Sequenz der SEQ ID NO.: 10 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 10 ist, die Sequenz der SEQ ID NO.: 11 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 11 ist, die Sequenz der SEQ ID NO.: 12 oder eine Sequenz, die zu mindestens 60% identisch mit der Sequenz der SEQ ID NO.: 12 ist, die Sequenz der SEQ ID NO.: 13 oder eine Sequenz, die zu mindestens 60% identisch mit der Sequenz der SEQ ID NO.: 13 ist, die Sequenz der SEQ ID NO.: 14 oder eine Sequenz, die zu mindestens 75% identisch mit der Sequenz der SEQ ID NO.: 14 ist, die Sequenz der SEQ ID NO.: 15 oder eine Sequenz, die zu mindestens 95% identisch mit der Sequenz der SEQ ID NO.: 15 ist, die Sequenz der SEQ ID NO.: 16 oder eine Sequenz, die zu mindestens 95% identisch mit der Sequenz der SEQ ID NO.: 16 ist, die Sequenz der SEQ ID NO.: 17 oder eine Sequenz, die zu mindestens 75% identisch mit der Sequenz der SEQ ID NO.: 17 ist, die Sequenz der SEQ ID NO.: 18 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 18 ist, die Sequenz der SEQ ID NO.: 19 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 19 ist, die Sequenz der SEQ ID NO.: 20 oder eine Sequenz, die zu mindestens 60% identisch mit der Sequenz

der SEQ ID NO.: 20 ist, die Sequenz der SEQ ID NO.: 21 oder eine Sequenz, die zu mindestens 90% identisch mit der Sequenz der SEQ ID NO.: 21 ist, die Sequenz der SEQ ID NO.: 22 oder eine Sequenz, die zu mindestens 70% identisch mit der Sequenz der SEQ ID NO.: 22 ist, die Sequenz der SEQ ID NO.: 23 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 23 ist, die Sequenz der SEQ ID NO.: 24 oder eine Sequenz, die zu mindestens 65% identisch mit der Sequenz der SEQ ID NO.: 24 ist, die Sequenz der SEQ ID NO.: 25 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 25 ist, die Sequenz der SEQ ID NO.: 26 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 26 ist, die Sequenz der SEQ ID NO.: 27 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 27 ist, die Sequenz der SEQ ID NO.: 28 oder eine Sequenz, die zu mindestens 75% identisch mit der Sequenz der SEQ ID NO.: 28 ist, die Sequenz der SEQ ID NO.: 29 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 29 ist, die Sequenz der SEQ ID NO.: 30 oder eine Sequenz, die zu mindestens 60% identisch mit der Sequenz der SEQ ID NO.: 30 ist, die Sequenz der SEQ ID NO.: 31 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 31 ist, die Sequenz der SEQ ID NO.: 32 oder eine Sequenz, die zu mindestens 55% identisch mit der Sequenz der SEQ ID NO.: 32 ist, die Sequenz der SEQ ID NO.: 33 oder die Sequenz der SEQ ID NO.: 34.

2. Nukleinsäuremolekül, das das Polypeptid nach Anspruch 1 kodiert.
3. Nukleinsäuremolekül, das komplementär zu dem Nukleinsäuremolekül nach Anspruch 2 ist.
4. Vektor, der das Nukleinsäuremolekül nach Anspruch 2 oder 3 enthält.
5. Nicht-menschlicher Wirt, der Polypeptid nach Anspruch 1 oder das Nukleinsäuremolekül nach Anspruch 2 oder 3 oder den Vektor nach Anspruch 4 enthält.

6. Wirt nach Anspruch 5, der eine Zelle ist.
7. Wirt nach 5, der ein transgenes nicht-menschliches Tier ist.
8. Wirt nach einem der Ansprüche 5 bis 7, der eine weitere, für die Cofaktor-Regenerierung geeignete Dehydrogenase oder ein diese Dehydrogenase kodierendes Nukleinsäuremolekül aufweist.
9. Wirt nach Anspruch 8, wobei die für die Cofaktor-Regenerierung geeignete Dehydrogenase eine Formiatdehydrogenase oder eine Glukosedehydrogenase ist.
10. Reaktionssystem enthaltend eine ein Substrat einer Dehydrogenase darstellende organische Verbindung, das Polypeptid nach Anspruch 1, den Vektor nach Anspruch 4 oder den Wirt nach einem der Ansprüche 5 bis 9 und gegebenenfalls einen Cofaktor für das Polypeptid nach Anspruch 1.
11. Reaktionssystem nach Anspruch 10, wobei die ein Substrat einer Dehydrogenase darstellende organische Verbindung eine Carbonylverbindung ist.
12. Reaktionssystem nach Anspruch 11, wobei die Carbonylverbindung eine Aldehyd oder ein Keton ist.
13. Reaktionssystem nach Anspruch 12, wobei das Keton ein unsymmetrisch substituiertes Keton ist.
14. Reaktionssystem nach Anspruch 10, wobei die ein Substrat einer Dehydrogenase darstellende organische Verbindung ein Alkohol ist.
15. Reaktionssystem nach Anspruch 14 wobei der Alkohol ein primärer Alkohol oder ein chiraler sekundärer Alkohol ist.

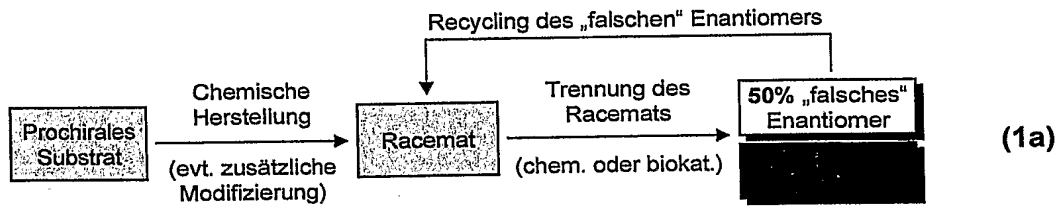
16. Reaktionssystem nach einem der Ansprüche 10 bis 15, wobei der Cofaktor NADH, NADPH, NAD^+ oder NADP^+ ist.
17. Verfahren zur Herstellung des Polypeptids nach Anspruch 1 oder eines durch das Nukleinsäuremolekül nach Anspruch 2 kodierten Polypeptids, wobei man den Wirt nach einem der Ansprüche 5 bis 9 züchtet und das Polypeptid isoliert.
18. Verfahren zur Herstellung des Polypeptids nach Anspruch 1 oder eines durch das Nukleinsäuremolekül nach Anspruch 2 kodierten Polypeptids, wobei man das Polypeptid aus einer Körperflüssigkeit oder Gewebeprobe des Wirts nach einem der Ansprüche 7 bis 9 isoliert.
19. Verfahren zur Herstellung einer ein Produkt einer Dehydrogenase darstellenden organischen Verbindung, wobei man eine ein Substrat einer Dehydrogenase darstellende organische Verbindung mit dem Polypeptid nach Anspruch 1, dem Wirt nach einem der Ansprüche 5 bis 9 oder mittels des Reaktionssystems nach einem der Ansprüche 10 bis 16 umsetzt.
20. Verfahren nach Anspruch 19, das ferner den Schritt der Isolierung des Produktes der Umsetzung umfasst.
21. Verfahren nach Anspruch 20, das ferner die Weiterverarbeitung des Produktes in ein Arzneimittel umfasst.
22. Verfahren nach Anspruch 20, weiter umfassend den Schritt der Weiterverarbeitung des Produkt in ein Folgeprodukt.
23. Verfahren nach Anspruch 22, weiter umfassend den Schritt der Formulierung des Folgeproduktes in der Herstellung eines Arzneimittels.
24. Verfahren nach einem der Ansprüche 19 bis 23, wobei das Produkt ein enantiomerreiner Alkohol ist.

25. Ligand, der das Polypeptid nach Anspruch 1 spezifisch bindet, wobei der Ligand kein Substrat oder Kofaktor des Polypeptids und kein davon erzeugtes Produkt ist.
26. Ligand nach Anspruch 25, der ein Antikörper oder ein Fragment oder Derivat davon, ein Aptamer, oder eine niedermolekulare Substanz ist.
27. Primer mit einer in Tabelle 1 dargestellten Sequenz.
28. Primerpaar mit in Tabelle 1 dargestellten Sequenzen, wobei der erste Primer des Primerpaares als Vorwärts-Primer und der zweite Primer des Primerpaares als Rückwärts-Primer zur Amplifizierung einer DNA-Sequenz dient.
29. Kit enthaltend
 - (a) das Polypeptid nach Anspruch 1;
 - (b) das Nukleinsäuremolekül nach Anspruch 2 oder 3;
 - (c) den Vektor nach Anspruch 4;
 - (d) den Wirt nach einem der Ansprüche 6 bis 9;
 - (e) das Reaktionssystem nach einem der Ansprüche 10 bis 16;
 - (f) den Liganden nach Anspruch 25 oder 26;
 - (g) mindestens einen Primer nach Anspruch 27; und/oder
 - (h) mindestens ein Primerpaar nach Anspruch 28

Zusammenfassung

Die Erfindung betrifft neue Polypeptide mit der biologischen Aktivität einer NAD- oder NADP-abhängigen Alkoholdehydrogenase. Des weiteren betrifft die Erfindung diese Polypeptide kodierende Nukleinsäuren, nicht-menschliche Wirte oder Wirtszellen sowie Reaktionssysteme, mit welchen gewünschte Produkte hergestellt werden können. Die erfindungsgemäßen Polypeptide werden bevorzugt in der Herstellung, ausgehend von Aldehyden bzw. Ketonen, von primären und enantiomerreinen, sekundären Alkoholen eingesetzt, die als Zwischenprodukte für Arzneimittel dienen können. Alternativ können die erfindungsgemäßen Polypeptide auch für die umgekehrte Reaktion, also die Oxidation von Alkoholen unter Ausbildung von Aldehyden bzw. Ketonen, eingesetzt werden.

- **Stand der Technik *via* Racematspaltung: mindestens 3-4 Schritte**



- **Biokatalytisches und nachhaltiges Konzept: Asymmetrische Biokatalyse**

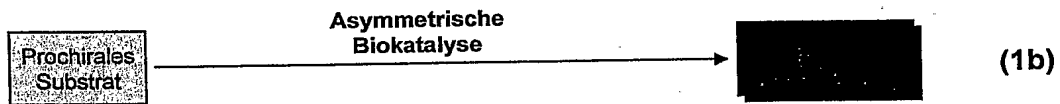


Fig. 1.

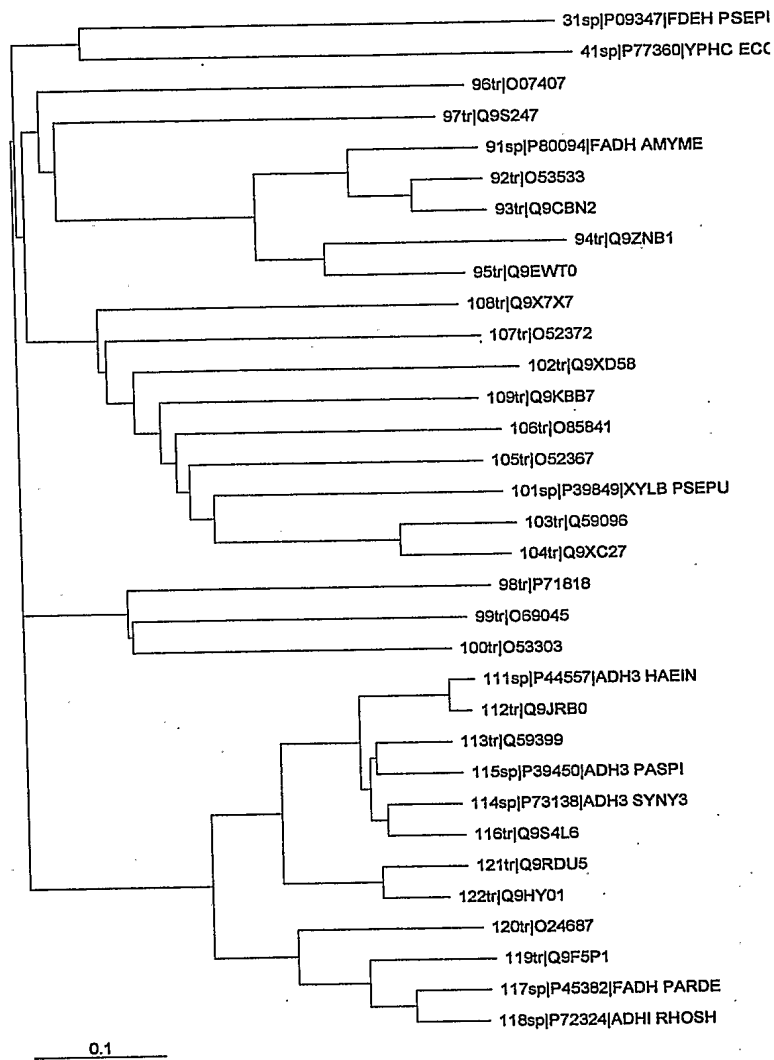


Fig. 2. Übersicht des Clusters 2 (= Primergruppe 2), basierend auf 33 Sequenzen

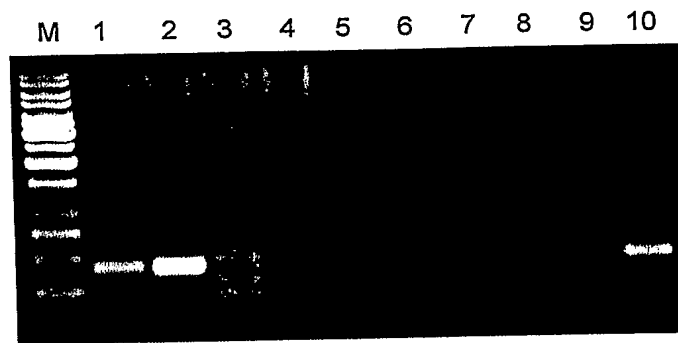


Fig. 3. PCR-Typisierung mit Primergruppe 2 unter Verwendung verschiedener Pools
 Spurenbelegung: M: Marker 1 kb –DNA-Leiter; Spur 1: Pool 1 mit Primer ADHM9 + 10;
 Spur 2: Pool 1 mit Primer ADHM11 + 12; Spur 3: Pool 1 mit Primer ADHM13 + 14; Spur 4:
 Pool 1 mit Primer ADHM15 + 16; Spur 5: Pool 2 mit Primer ADHM 9 + 10; Spur 6: Pool 2
 mit Primer ADHM11 + 12; Spur 7: Pool 2 mit Primer ADHM13 + 14; Spur 8: Pool 2 mit
 Primer ADHM15 + 16; Spur 9: Pool 3 mit Primer ADHM9 + 10; Spur 10: Pool 3 mit Primer
 ADHM11 + 12